National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

NL-339 (r. 88/04) c

Canada

UNIVERSITY OF ALBERTA

LONG-RANGE ADAPTIVE PREDICTIVE CONTROL
EXTENSIONS AND APPLICATIONS

BY

KUN-YU KWOK ©

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE
OF
DOCTOR OF PHILOSOPHY
IN
PROCESS CONTROL

DEPARTMENT OF CHEMICAL ENGINEERING

EDMONTON, ALBERTA
SPRING, 1992

National Library of Canada
Canadian Theses Service
Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada
Service des thèses canadiennes

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-73090-0

Canada

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Kun-Yu Kwok

TITLE OF THESIS: Long-Range Adaptive Predictive Control: Extensions and Applications

DEGREE: Doctor of Philosophy

YEAR THIS DEGREE GRANTED: 1992

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

ADDRESS:

11245-78 Avenue
Edmonton, Alberta
Canada T6G 0M8

Date: April 1, 1992

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled

Long-Range Adaptive Predictive Control: Extensions and Applications

submitted by Kun-Yu Kwok in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Process Control.

....
Dr. S.L. Shah
(Supervisor)

....
Dr. A.S. Clanachan
(Pharmacology)

....
Dr. G.A. Dumont
(Electrical Engineering, UBC)

....
Dr. N.G. Durdle
(Electrical Engineering)

....
Dr. B.A. Finegan
(Anaesthesia)

....
Dr. D.G. Fisher
(Chemical Engineering)

Date: March 16, 1992

To My Parents

獻給摯愛的雙親
郭 斌 樑先生
郭張潔英女士

ABSTRACT

Long-range predictive control (LRPC) has been gaining acceptance in many industrial applications because of its unique ability to implement multi-step optimization. This thesis is concerned with the extension of the performance criteria to include a terminal matching condition in both the control and identification strategies, and the implementation of adaptive LRPC for the regulation of mean arterial pressure.

A weighting term on the square of prediction error at steady state was added as a terminal condition in the LRPC performance criteria. The properties of this new weighting term have been examined using the Generalized Predictive Control (GPC) algorithm of Clarke *et al.* (1987), a widely used version of the class of LRPC. Closed-loop analysis shows that the steady-state error weighting is not only comparable to large prediction horizons, but also has additional advantages over ordinary control weighting. Most interestingly, this weighting makes it possible to remove the additional closed-loop order introduced by the integral control action and yet does not result in any offset at steady state even in the presence of a gain mismatch. Two strategies are suggested for adjusting such a weighting term as a primary tuning parameter while other LRPC tuning parameters are kept constant.

The dual of the steady-state error weighting term in long-range predictive identification (LRPI) was examined in the context of the $L(q^{-1})$-filtering algorithm developed by Shook *et al.* (1991). For the prediction horizon $n2$ tending to infinity, Shook's algorithm causes the whole ARMAX modeling scheme to change from an equation error scheme to an output error one. However, this convergence property is

not valid for the ARIMAX model because $L(q^{-1})$ with ARIMAX does not converge to a finite polynomial. Although an extension of Shook's LRPI algorithm to include a terminal matching condition is not feasible, analyzing the derivation of LRPI algorithm does reveal the fact that the terminal condition can be indirectly realized by identification of the process gain. Thus, a simple algorithm similar to the non-minimal model predictor (Lu and Fisher, 1990) is proposed for on-line estimation of a process gain. As a result, an overall performance criterion for a multi-step, adaptive, predictive control including the terminal condition is realized by the combination of a long-range predictive control law such as GPC, the steady-state error weighting, Shook's (1991) LRPI algorithm, and an on-line gain estimation algorithm such as the one proposed in this thesis. The performance using this overall performance criterion was shown by simulation to be very effective and to yield results superior to classical GPC. Even with a very small prediction horizon and large model-plant mismatch, an accurate estimation of process gain combined with the steady-state error weighting improves the overall robustness of the controller.

The combination of GPC, LRPI, and steady-state error weighting was also tested experimentally on a pilot-scale continuously-stirred heating system and the regulation of mean arterial pressure (MAP). Because of the latter application, this thesis also describes the development of a closed-loop control system for automatic drug delivery. This truly model-based adaptive control system, a combination of GPC with steady-state error weighting and LRPI, accomplishes an overall performance objective function in which identification and control are mutually compatible. It was developed in a real-time, multitasking software environment on a personal computer

so that on-line changes are possible. The infusion of sodium nitroprusside (SNP), a clinically used vasodilator, is managed by a drug delivery pump controlled by the personal computer. All in-house programs were coded in "C"-language. Several experimental studies of the regulation of MAP using SNP infusion were conducted in an ethics-approved manner. The effectiveness of this control system was evaluated by a series of disturbances induced by the infusion and injection of norepinephrine, adenosine mono-phosphate, and SNP. It is concluded that the system performs favorably even on a highly non-linear and time-varying process in a delicate environment such as MAP regulation.

ACKNOWLEDGEMENT

This work would not have been completed without the help and support from a number of people in the past several years. In particular, I am grateful to Dr. Sirish Shah for his personal guidance and supervision, Dr. Clanachan for his constant encouragement, Dr. Finegan for his genuine enthusiasm, and Dr. Fisher for his critical proof-re~~e~~ading of my thesis.

My fellow graduate students also deserve many thanks: Mike Foley who gave me several lectures on adaptive control; Eric Lau, Pradyumna Namdev, Paul Nawrocki, Don Ryland, and Yan-sung Yiu who struggled with me through the course work; Shuk-yee Ho, Wei-ping Lu, and Dave Shook who shared with me the excitement of my research; and those who were in Process Control. I am also indebted to the DACS Center staff and Ireneusz Wachowski for providing technical assistance in preparing the experimental work.

Financial support from the Department of Chemical Engineering and the Natural Sciences and Engineering Research Council of Canada is much appreciated.

I would like to express my appreciation to several groups of individuals for their constant spiritual support through their prayers: my brother-in-law Stephen Yeung, now with the Lord; my sisters and brother Trudy, Fong, and Caleb; my buddies Kent Leung, Philip Pang, and Devonne Lui; and especially my ἀγαπητός Juliana Chi-yan Ching whose love, care, and support are always with me.

This thesis is respectfully dedicated to my parents because their love allowed me to experience the true ἀγάπη of my Lord Christ Jesus who has sacrificed Himself on the cross on Calvary for our sins.

Contents

1 Introduction 1
1.1 Scope and Objectives of Study 2
1.2 Thesis Organization 4
Bibliography 6

2 LRPC with a Terminal Matching Condition: Theory 8
2.1 Introduction 8
2.2 Process Modeling 12
2.2.1 Transfer Function Model with d.c. Gain 12
2.2.2 Transfer Function Model with Integrator 14
2.2.3 Convolution Model 20
2.3 Control Algorithm 23
2.3.1 Derivation 23
2.3.2 Control Law in Linear Form 26
2.4 Closed-Loop Analysis of γ-Weighting 28
2.4.1 Closed-Loop Transfer Function 28
2.4.2 Minimum Variance Control with γ-Weighting 31
2.4.3 Mean-Level Property 36
2.5 Conclusions 39
Bibliography 41

3 Evaluation of LRPC with a Terminal Matching Condition 44
3.1 Introduction 44
3.2 Control Algorithm Review 45
3.3 Performance Tuning of γ-Weighting 48
3.3.1 Strategy #1: Deadbeat Approach 48
3.3.2 Strategy #2: Mean-Level Approach 50
3.4 Demonstration of γ-Weighting 52
3.4.1 GPC with γ-Weighting 54
3.4.2 Adaptive GPC with γ-Weighting 66
3.5 Conclusions 76
Bibliography 77

4 **LRPI with a Terminal Matching Condition** **79**
4.1 Introduction 79
4.2 Overall Adaptive Control Objective 81
4.2.1 Finite Horizon Predictive Control 81
4.2.2 Finite Horizon Predictive Control with a Terminal Matching Condition 83
4.3 LRPI with a Terminal Matching Condition 85
4.3.1 LRPI Algorithm 85
4.3.2 Identification of Process Gain 87
4.4 Synopsis of an Overall Adaptive Controller 88
4.5 Simulation Studies 89
4.5.1 Open-Loop Identification 89
4.5.2 Closed-Loop Control 93
4.6 Conclusions 98
Bibliography 102

5 **Automated Blood Pressure Regulation: A Survey** **103**
5.1 Introduction 103
5.2 Sheppard's System 104
5.3 IVAC Titrator 109
5.4 Other Ventures 111
5.4.1 Single-Step Predictive Optimization 111
5.4.2 Multi-Step Predictive Optimization 116
5.4.3 Multiple Drug Infusion 117
5.4.4 Expert and Fuzzy Control System 119
5.5 Conclusions 120
Bibliography 122

6 **Development of LRPC System for Mean Arterial Pressure Regulation** **131**
6.1 Introduction 131
6.2 Controller Design 133
6.2.1 Model-Based Predictive Control 133
6.2.2 Constrained Optimization 136
6.2.3 Process Identification 137
6.2.4 Implementation 137
6.3 System Description 138
6.3.1 Equipment 138
6.3.2 Software 140
6.4 Experimental Studies 141
6.4.1 Methods 141
6.4.2 Results 143
6.5 Discussion 149

6.6 Conclusions 151
Bibliography 152

7 Evaluation of LRPC System for Mean Arterial Pressure Regulation 156
7.1 Introduction 156
7.2 Process Control Strategy 158
    7.2.1 Control Scheme 158
    7.2.2 Identification 159
7.3 System Description 160
    7.3.1 Equipment 160
    7.3.2 Control Program 161
7.4 Experimental Studies 161
    7.4.1 Methods 161
    7.4.2 Results 165
7.5 Discussion 168
7.6 Conclusions 173
Bibliography 175

8 Conclusions 178
8.1 General Discussions and Conclusions 178
8.2 Recommendations 182
Bibliography 183

A Example: Simultaneous Offset Removal and Integrator-Pole Elimination by Steady-State Error Weighting 185

List of Figures

2.1 Long-range predictive control with steady-state error weighting 10
2.2 A general closed-loop control system 29

3.1 Trajectories of γ for different values of $n2$ for the special case of $nu = 1$ 53
3.2 Closed-loop root locus of process A using minimum variance control with λ-weighting 55
3.3 Closed-loop root locus of process A using minimum variance control with γ-weighting 56
3.4 Closed-loop root locus of process A using dead-beat control with γ-weighting 58
3.5 Closed-loop performance for process A using dead-beat control with γ-weighting 59
3.6 Closed-loop poles of process B using standard GPC 60
3.7 Closed-loop root locus of process B using GPC with γ-weighting . . . 62
3.8 Closed-loop root locus of process B using GPC with λ-weighting . . . 63
3.9 Closed-loop performance for process B using GPC with γ-weighting . 64
3.10 Closed-loop performance for process B using GPC with λ-weighting . 65
3.11 Step response trajectories of modeling process C using various modeling horizons 67
3.12 Closed-loop performance for process C using GPC with γ-weighting . 68
3.13 Closed-loop performance for process C using adaptive GPC with γ-weighting 70
3.14 Schematic diagram of continuous stirred-tank heater 71
3.15 Closed-loop performance of continuous stirred-tank heater for ($n1$, $n2$, nu, γ)=(1,5,1,.6) 73
3.16 Closed-loop performance of continuous stirred-tank heater for ($n1$, $n2$, nu, γ)=(1,5,1,0) 74
3.17 Closed-loop performance of continuous stirred-tank heater for ($n1$, $n2$, nu, γ)=(1,10,1,0) 75

4.1 Schematic diagram of an adaptive controller using LRPI/LRPC with terminal matching conditions 90
4.2 Comparison of step response trajectories between Rohrs' model and three LRPI models 92

4.3 Closed-loop response using GPC and LRPI, $n2 = 2$ for $t = 1$ to 250, $n2 = 5$ for $t = 251$ to 450 94
4.4 Closed-loop response using GPC and LRPI, $n2 = 5$ for $t = 1$ to 250, $n2 = 10$ for $t = 251$ to 450 95
4.5 Closed-loop response using GPC plus steady-state error weighting and LRPI plus gain estimation, $n2 = 2$, $\gamma = .05, .05, .1, .5, 1$ for each setpoint change 96
4.6 Closed-loop response using GPC plus steady-state error weighting and LRPI plus gain estimation, $n2 = 5$, $\gamma = .001, .001, .01, .1, .5$ for each setpoint change 97
4.7 Model gain trajectory corresponding to fig. 4.3 99
4.8 Model gain trajectory corresponding to fig. 4.5 100

5.1 The schematic diagram of a control system by Slate and Sheppard . . 107

6.1 Block diagram of the adaptive blood pressure control system 139
6.2 Block digram of the MULTICON real-time, multi-tasking control program 142
6.3 Open-loop step responses of MAP 144
6.4 Trajectory of MAP in response to disturbances under unconstrained GPC control 145
6.5 Trajectory of MAP in response to disturbances at open-loop fixed SNP infusion rate 147
6.6 Trajectory of MAP in response to disturbances under constrained GPC control 148

7.1 Schematic diagram of the experimental setup 163
7.2 Average deviation of MAP from the target during setpoint change . . 166
7.3 Average deviation of MAP from the target during NE infusion 167
7.4 Average deviation of MAP from the target during SNP infusion . . . 169
7.5 Average deviation of MAP from the target during NE injection . . . 170
7.6 Average deviation of MAP from the target during SNP injection . . . 171

List of Tables

3.1 Process models for computer simulations 54
3.2 Effect of λ-weighting on Process B 61
3.3 Tuning parameter configurations 72

6.1 Challenge procedure during closed-loop control 146

List of Symbols

$A(q^{-1})$ = process model denominator polynomial
$\bar{A}(q^{-1})$ = product of $A(q^{-1})$ and Δ
a_i = i^{th} coefficient of $A(q^{-1})$
a_e = approximation error in convolution model
a_{ss} = reference value in convolution model
$B(q^{-1})$ = process model numerator polynomial
b_i = i^{th} coefficient of $B(q^{-1})$
b'_i = i^{th} coefficient of the numerator polynomial in G_p
$C(q^{-1})$ = noise model polynomial
d = d.c. gain in process model
$E_j(q^{-1})$ = quotient polynomial from j^{th}-step ahead Diophantine identity $\frac{T}{A\Delta}$
$E'_j(q^{-1})$ = quotient polynomial from j^{th}-step ahead Diophantine identity $\frac{T}{A}$
e_i = i^{th} coefficient of $E_j(q^{-1})$
$F_j(q^{-1})$ = remainder polynomial from j^{th}-step ahead Diophantine identity $\frac{T}{A\Delta}$
$F'_j(q^{-1})$ = remainder polynomial from j^{th}-step ahead Diophantine identity $\frac{T}{A}$
$f_i(q)$ = i^{th} first-order fraction of $\frac{1}{A}$
$f(t+i)$ = i^{th} open-loop process output prediction assuming no future control action
$\mathbf{f}$ = vector of $f(t+i)$ where i is from $n1$ to $n2$
$\mathbf{f_s}$ = vector of $f(s)$
$\mathbf{G}$ = matrix of step response elements
$\mathbf{G_s}$ = lower triangular matrix containing only g_s's
$G_j(q^{-1})$ = quotient polynomial from j^{th}-step ahead Diophantine identity $\frac{E_jB}{T}$
$G'_j(q^{-1})$ = quotient polynomial from j^{th}-step ahead Diophantine identity $\frac{E'_jB}{T}$
$G_p(q^{-1})$ = transfer function of any process
g_i = i^{th} coefficient of $G_j(q^{-1})$
$H_j(q^{-1})$ = remainder polynomial from j^{th}-step ahead Diophantine identity $\frac{E_jB}{A}$
$H'_j(q^{-1})$ = remainder polynomial from j^{th}-step ahead Diophantine identity $\frac{E'_jB}{A}$
h_i = i^{th} impulse response coefficient
$h_{s,i}$ = i^{th} coefficient of $H_s(q^{-1})$
$I(s)$ = nitroprusside infusion rate in continuous time domain
J = performance index
k = time delay in number of sampling intervals

N = state-space model order
n = number of step response coefficients
$n1$ = minimum output prediction horizon
$n2$ = maximum output prediction horizon
na = order of $A(q^{-1})$
np = number of predictions in the prediction horizon ($n2$-$n1+1$)
ns = maximum output prediction horizon less than $n2$
$\bar{na}$ = order of $\bar{A}(q^{-1})$
nb = order of $B(q^{-1})$
nt = order of $T(q^{-1})$
nu = control horizon
$P(s)$ = mean arterial pressure in continuous time domain
p = first-order coefficient in the denominator of G_p
q^{-1} = backward shift operator
$R(q^{-1})$ = polynomial in the linear form of GPC control law
r = d.c. gain at steady state
r_i = i^{th} root of $A(q^{-1})$
$S(q^{-1})$ = polynomial in the linear form of GPC control law
s = value at steady state
s_i = multiplicity of r_i
$T(q^{-1})$ = observer design polynomial
T = model horizon for convolution model, or current time in J
t_i = i^{th} coefficient of $T(q^{-1})$
$\mathbf{U}$ = vector of incremental control inputs
$u(t)$ = control input
$V(q^{-1})$ = polynomial in the linear form of GPC control law
$\mathbf{W}$ = vector of setpoint trajectory
$w(t)$ = setpoint
$x(t)$ = noise input
$\mathbf{Y}$ = vector of output prediction values
$y(t)$ = process output
z^{-1} = z-transform backward shift operator

Greek symbols

α = recirculation fraction
α_i = i^{th} element of the first row of $[\mathbf{G}^T\mathbf{G} + \Lambda + \mathbf{G_s}^T\Gamma\mathbf{G_s}]^{-1}\mathbf{G}^T$
α_s = sum of β_i from $n1$ to $n2$
α_u = rate constraint limit for constrained GPC algorithm
β_i = i^{th} element of the first row of $[\mathbf{G}^T\mathbf{G} + \Lambda + \mathbf{G_s}^T\Gamma\mathbf{G_s}]^{-1}\mathbf{G_s}^T\Gamma$
Γ = diagonal matrix containing control steady-state weightings
γ = control steady-state weighting

γ_m = a critical steady-state weighting which makes closed-loop control stable
Δ = differencing operator, $1 - q^{-1}$
δ = degree of polynomial
δ_i = i^{th} coefficient of $\mathcal{D}$
Λ = diagonal matrix containing control weightings
λ = control weighting
$\xi(t)$ = uncorrelated random sequence of zero mean
ρ_i = amplitude constraint limits for constrained GPC algorithm
σ_i = gain of $f_i(q)$
τ = time constant
τ_s = sampling time
ω = frequency domain

Superscripts

$\hat{}$ = estimated value
$*$ = summation of all impulse response coefficients multiplied by the corresponding control actions
0 = denote true model polynomial
T = transposition

Subscripts

C = control
$FHPC$ = finite horizon predictive control
f = signal filtered by $\frac{1}{T(q^{-1})}$
GPC = generalized predictive control
ID = identification
LLS = least squares with $L(q^{-1})$-filter
$LRPC$ = long-range predictive control
$LRPI$ = long-range predictive identification
LS = least squares
MV = minimum variance control
S = steady state
X = cross-product term

Others

$\mathcal{C}$ = characteristic equation
$\mathcal{D}$ = a portion of $\mathcal{C}_s$

Chapter 1

Introduction

Closed-loop proportional, integral, and derivative (PID) control has been a popular control mechanism in chemical, mining, metallurgical and the pulp-and-paper industries. Early classical feedback control methods have been progressively supplemented by other advanced predictive and adaptive techniques since the landmark papers by Åström and Wittenmark (1973), Clarke and Gawthrop (1975), etc. More recently, the successful applications of Model Algorithmic Control (Richalet *et al.*, 1978) and Dynamic Matrix Control (Cutler and Ramaker, 1980) have been followed by several unifying approaches such as Internal Model Control (Garcia and Morari, 1982) and Generalized Predictive Control (GPC) (Clarke *et al.*, 1987). As these advanced control techniques receive widespread acceptance among chemical industries (Seborg *et al.*, 1986; Lambert, 1987; M'Saad *et al.*, 1987) , applications in other areas are also being pursued.

One of the areas where closed-loop control has currently received much attention is in the control of biomedical systems (Vozeh and Steimer, 1985; Linkens, 1984; Linkens and Hacisalihzade, 1990). The human body from a physiological point of view can be thought of as being a chemical plant in which numerous chemical reactions and mass transfer operations occur. The brain serves as the main control center which requires the autonomic or central nervous system to provide voluntary as well as involuntary feedback control signals. When the mini-"chemical plant" fails to maintain one or more physiological variables within certain desirable limits, the administration of therapeutic drugs becomes necessary. If the control of the physiological variable requires continuous drug intervention, a closed-loop monitoring system with automated drug delivery will potentially improve the quality of control as well as simplifying the labor-intensive drug administration procedure currently practiced by experienced medical personnel. This incentive has brought a number of investigations into automated drug delivery systems especially for mean arterial pressure (MAP) regulation because most cardiac patients require intraoperative and postoperative drug therapy for induced hypotension.

1.1 Scope and Objectives of Study

An early closed-loop control system designed for MAP regulation was based on the classical proportional-integral-derivative (PID) feedback method (Sheppard *et al.*, 1975). The manipulated variable for induced hypotension was the infusion of sodium nitroprusside (SNP), a commonly used vasodilator. Because a PID-based algorithm was used, frequent tuning was necessary to maintain desired performance (Sheppard, 1981). Since body sensitivity and the dynamic response of MAP to SNP infusion vary from patient to patient or even during the course of infusion on the same patient, advanced control techniques with some degree of adaptivity should be more beneficial (Katona, 1982). Although several adaptive and predictive control strategies have been investigated (see Chapter 5), their control objectives are all based on single-step ahead, single-point optimization strategies that are highly sensitive to model-plant mismatch and varying time delay.

Long-range predictive control (LRPC) which considers the minimization of the sum of prediction errors over a certain time horizon appears to be a better alternative. One popular version of LRPC is commonly known as GPC. With the long-range predictive identification (LRPI) algorithm developed by Shook *et al.* (1991), a complete overall performance criterion is "optimized". However, the long-range prediction requirement necessarily makes the computation load of an adaptive controller very heavy for each new control action to be calculated. On the other hand, a relatively short output prediction horizon may result in overly strong control action. The long-range control algorithm would be less computationally-intensive if a shorter output prediction horizon were used with a facility for retaining the long-range property by considering the output prediction error at infinite time. Therefore, the scope of this work is three-fold: first, to include a terminal matching condition in the LRPC formulation; second, to develop the theoretical results in design procedure; and third, to implement a robust adaptive closed-loop control system for MAP regulation. One specific objective of this research is to develop a closed-loop biomedical controller which makes use of LRPC algorithm to achieve high quality and robust closed-loop control of arterial blood pressure.

In summary, the objectives of this thesis are:

- explore the potential and properties of a terminal matching condition in the context of LRPC,
- obtain a control-relevant identification strategy for LRPI controllers using the terminal matching condition,
- develop a long-range, adaptive control system for automatic drug delivery,
- experimentally evaluate the performance of the controller in the regulation of MAP.

1.2 Thesis Organization

The first part of this thesis from Chapters 2 to 4 concentrates on the development of relevant theory. The second part (Chapter 5 onwards) deals with the literature review, implementation, and evaluation of a closed-loop drug delivery system.

Chapter 2 describes the formulation and closed-loop analysis of the terminal matching condition as an extension to the basic LRPC objective. The terminal condition is formulated as a weighting on the output steady-state error. This chapter also shows that the steady-state error weighting is beneficial under both parametric and convolution modeling techniques.

Since this new weighting term becomes an additional tuning parameter, two tuning strategies are proposed in Chapter 3. An evaluation of this new tuning parameter compared with other tuning parameters is also included in this chapter.

LRPI with the proposed terminal condition is discussed in Chapter 4. It is found that a combination of the LRPI algorithm developed by Shook *et al.* (1991) and an adaptive gain estimation algorithm results in a LRPC-relevant identification strategy. A similar version of the non-minimal predictive control of Lu and Fisher (1990) is suggested as a gain estimation algorithm .

Chapter 5 serves as a bridge between control theory and biomedical application. It contains a critical and thorough review of the current development in automated blood pressure regulation. Sheppard's MAP controller (Slate *et al.*, 1980) and the IVAC Titrator (Cosgrove III *et al.*, 1989) are surveyed first, followed by other developments such as predictive control strategies and expert control systems.

The design and development of a long-range, adaptive, and predictive control system for blood pressure regulation are described in Chapter 6. The system features both constrained and unconstrained GPC, and has been tested in a preliminary study

(Kwok *et al.*, 1991).

Chapter 7 systematically evaluates the final version of the control system which has been implemented with the steady-state error weighting term and LRPI.

Each chapter represents an individual effort and thus is furnished with a more detailed introductory section. Since a majority of the chapters have been either published or submitted for publication, the thesis is arranged in a paper format acceptable to the Faculty of Graduate Studies and Research. Overall conclusions and recommendations for future research are found in the final chapter of the thesis.

Bibliography

Clarke, D.W., C. Mohtadi, and P.S. Tuffs. Generalised predictive control - Part I. The basic algorithm. *Automatica*, 23(2):137–148, 1987.

Cosgrove III, D.M., J.H. Petre, J.L. Waller, J.V. Roth, C. Shepherd, and L.H. Cohn. Automated control of postoperative hypertension: A prospective, randomized multicenter study. *Ann. Thorac. Surg.*, 47:678–683, 1989.

Cutler, C.R. and B.L. Ramaker. Dynamic matrix control — a computer control algorithm. In *Proc. JACC*, San Francisco, 1980.

Garcia, C.E. and M. Morari. Internal model control. 1. A unifying review and some new results. *Ind. Eng. Chem. Process Des. Dev.*, 21:308–323, 1982.

Katona, P.G. Automated control of physiological variables and clinical therapy. *CRC Critical Reviews in Biomedical Engineering*, 8(4):281–310, 1982.

Kwok, K.Y., R.K. Mutha, S.L. Shah, A.S. Clanachan, and B.A. Finegan. Constrained long-range adaptive predictive control of arterial blood pressure. *International Journal of Adaptive Control and Signal Processing*, 5(6):363–374, 1991.

Lambert, E.P. *Process Control Applications of Long Range Prediction.* D.Phil. thesis, University of Oxford, 1987.

Linkens, D.A. Computer control in biomedicine. In S. Bennett and D.A. Linkens, editors, *Real-Time Computer Control*, chapter 14. P. Peregrinis, 1984.

Linkens, D.A. and S.S. Hacisalihzade. Computer control systems and pharmalogical drug administration: a survey. *J. Med. Eng. & Tech.*, 14(2):41–54, 1990.

Lu, W. and D.G. Fisher. Nonminimal model based long range predictive control. In

Proc. American Control Conference, pages 1607–1613, San Diego, CA, 1990.

M'Saad, M., M. Duque, and E. Irving. Thermal process robust adaptive control: an experimental evaluation. In *Proc. 10th IFAC World Congress*, Munich, 1987.

Richalet, J., A. Rault, J.L. Testud, and J. Papon. Model predictive heuristic control: applications to industrial processes. *Automatica*, 14:413–428, 1978.

Seborg, D.E., T.F. Edgar, and S.L. Shah. Adaptive control strategies for process control: a survey. *A.I.Ch.E. Journal*, 32:881–913, 1986.

Sheppard, L.C. Computer control of the infusion of vasoactive drugs. *Ann. Biomed. Eng.*, 8:431–444, 1981.

Sheppard, L.C., N.T. Kouchoukos, J.F. Shotts, and F.D. Wallace. Regulation of mean arterial pressure by computer control of vasoactive agents in postoperative patients. In *Computers in Cardiology*, pages 91–94, Rotterdam, Netherlands, October 2–4 1975.

Shook, D.S., C. Mohtadi, and S.L. Shah. Identification for long range predictive control. *IEE Proc.-D*, 138(1):75–84, 1991.

Slate, J.B., L.C. Sheppard, V.C. Rideout, and E.H. Blackstone. Closed-loop nitroprusside infusion: modeling and control theory for clinical application. In *IEEE 1980 International Symposium on Circuits and Systems*, pages 482–488, 1980.

Vozeh, S. and J.L. Steimer. Feedback control methods for drug dosage optimisation. *Clinical Pharmacokinetics*, 10:457–476, 1985.

Chapter 2

LRPC with a Terminal Matching Condition: Theory

2.1 Introduction

The early methods for minimum variance (Åström and Wittenmark, 1973) and generalized minimum variance control (Clarke and Gawthrop, 1975) were aimed at providing tight control of processes by minimizing the variance of the process output predictions. A key factor in the successful use of these methods is correct *prior* selections of the model order, n, and time delay, k, such that accurate k-step ahead predictions are made. Poor selection of k or a varying time delay affects the control performance or even leads to instability. Improper selections of model order may result in similar problems although over-specified model order may not be as sensitive as the uncertainties in time-delay. Recent developments in long-range predictive control (LRPC) which minimize the sum of prediction errors over a user-specified time horizon provide an alternate solution to the above mentioned drawbacks (Richalet *et al.*, 1978; Cutler and Ramaker, 1980; Clarke and Mohtadi, 1989). Since the strength of the LRPC algorithm lies in the long-range predictions of process outputs and control actions, these two predictions which have a direct influence on stability and dynamic response in any long-range predictive controller (Garcia and Morari, 1982) naturally become the tuning parameters. The choices of prediction horizons are less influential than that of the time-delay or model order. However, the long-range predictive nature necessarily makes the computation load of an adaptive controller very heavy for each new control action to be calculated. On the other hand, a relatively short output prediction horizon may result in vigorous control actions or even unstable control. Moreover, the long-range control algorithms such as MAC (Richalet *et al.*, 1978), DMC (Cutler and Ramaker, 1980), and MOCCA (Sripada and Fisher, 1985) which employ a convolution model as the internal model (Garcia *et al.*, 1989) to encompass most of the process dynamics require as many as 50 step or impulse response coefficients, a size not economically sound for adaptive control. This paper proposes a modified approach to the LRPC scheme which introduces an output terminal condition in the minimization of a multi-step prediction cost function. The terminal condition is applied in the form of weighting on the square of the error between the predicted process steady-state value and the setpoint. This weighting combined with a relatively short output prediction horizon is able to resolve the trade off between economic computation and long-range robustness. The concept of steady-state error weighting is depicted in fig. 2.1. An ordinary LRPC algorithm considers process output predictions over a large horizon from $n1$ to $n2$. The modified approach would consider a relatively shorter horizon from $n1$ to ns plus the term at steady-state so as to ensure the terminal matching of process output to the setpoint.

Steady-state information has been very useful in many control applications. For example, steady-state gains have been routinely used in steady-state optimization of chemical processes (McFarlane and Bacon, 1989). In some cases, steady-state models from non-linear modeling techniques were also used to determine the current optimum and the direction of movement of the manipulated variable (Bamberger and

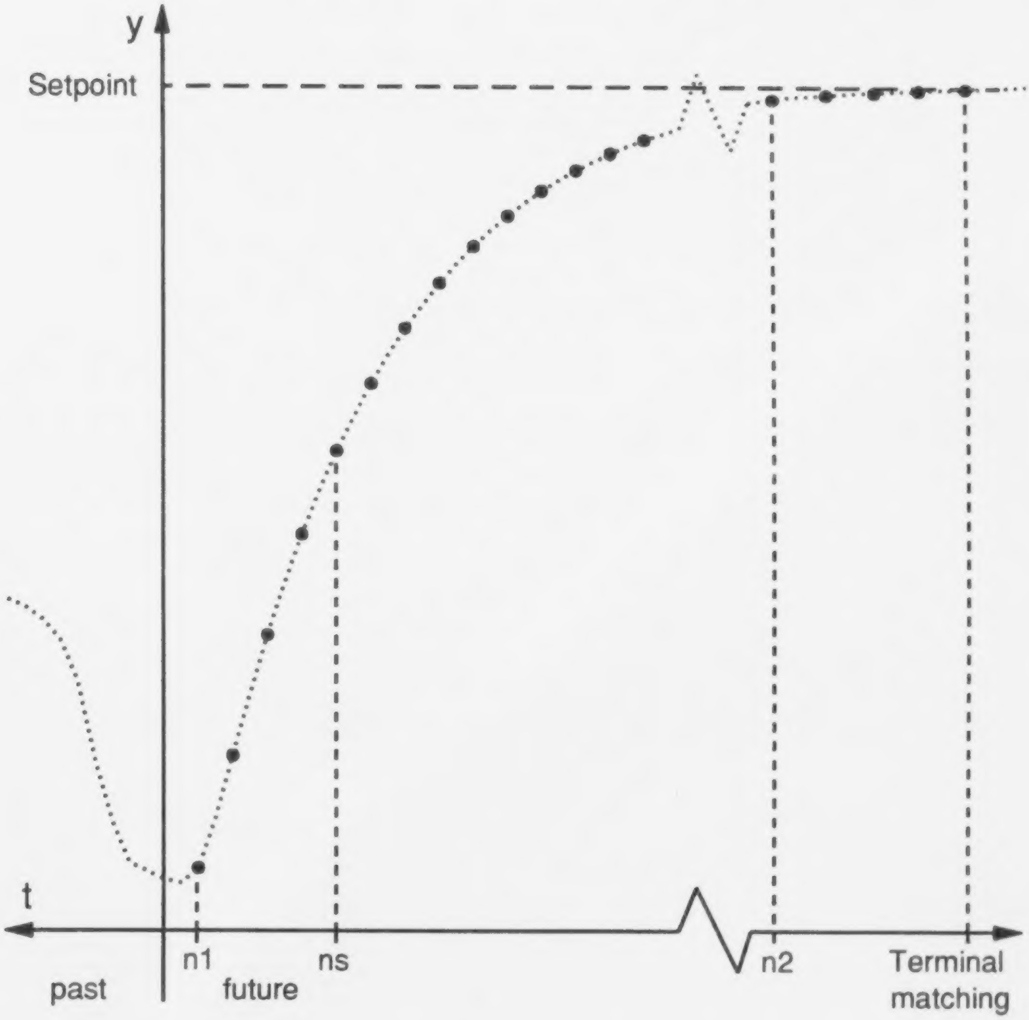
y
Setpoint
t
n1
ns
n2
Terminal matching
past
future

Figure 2.1: Long-range predictive control with steady-state error weighting

Isermann, 1978); then, the manipulated variable was moved in the direction of the steepest descent to minimize the objective function (Garcia and Morari, 1981; Garcia and Morari, 1984; Bhattacharya and Joseph, 1982). But never has the weighted steady-state value been used as a tuning parameter. Papadoulis *et al.* (1987) have developed a cautious self-tuning controller comprising a minimum variance part and a cautious part. The cautious part was implemented by penalizing in the control law the square of control deviation from a constant that approximates the control action at the target steady state. However, the steady-state value was not estimated on-line. In addition, it has little direct influence on dynamic response because the contribution of the cautious part diminishes as identification improves.

The contributions of this study are:

1. The incorporation of a weighted steady-state term as a tuning parameter in the performance criteria of LRPC using both transfer function and convolution models. The incorporation of such weighting in the optimization criterion by itself does not reduce the computation load. In fact, the use of conventional prediction techniques with steady-state weighing would still require just as many iterations of the Diophantine identity or step-response coefficients to predict the future steady-state value of the output. Technical results presented in Section 2.2 provide an alternative to this difficulty so that fewer iterations are required because of using a smaller prediction horizon with transfer function models. With a convolution model, a complete set of step response coefficients is not required. Only a few dynamics response coefficients plus the steady-state gain are required in deriving a long-range predictive controller. The latter idea relies on exact representation of the high-frequency dynamics of the process by the first few step-response coefficients and approximating the remainder terms

(usually low frequency terms) by a first-order model.

2. Development of a closed-loop long-range predictive controller which possesses a finite-time horizon with a weighting on the infinite-time horizon for terminal matching.

3. Analysis of the steady-state error weighting (γ-weighting) in the context of its effect on closed-loop stability and performance.

2.2 Process Modeling

Most model predictive control techniques are based on the optimization of a quadratic objective function involving the error between the setpoint and the predicted outputs. The distinguishing feature among different techniques is the model type and structure for making the predictions. The most commonly used model types are: discrete-time transfer functions describing the input-output relationship, or discrete-time convolution models in the form of step or impulse response coefficients. The use of these models for finite-horizon plus steady-state prediction is considered in the following sections.

2.2.1 Transfer Function Model with d.c. Gain

A single-input, single-output system can be described by a time-series type autoregressive moving-average model with an exogenous input(ARMAX),

$$A(q^{-1})y(t) = q^{-k}B(q^{-1})u(t) + d + C(q^{-1})\xi(t) \tag{2.1}$$

where $A(q^{-1})$, $B(q^{-1})$, and $C(q^{-1})$ are polynomials in the backward shift operator q^{-1} such that

$$q^{-1}y(t) = y(t-1)$$

$\xi(t)$ is an uncorrelated random noise sequence with zero mean. k is the process time delay. d represents the non-zero process output when the control is steady at zero. The order of the system is defined by the degree of the polynomial $A(q^{-1})$.

Since the parametric polynomials $A(q^{-1})$ and $B(q^{-1})$ for the true system are unknown, their estimates from recursive identification are used for controller design according to the certainty equivalence principle. The noise polynomial $C(q^{-1})$ which describes "colored" disturbances on the output is either estimated on-line by the Extended Least Squares identification method or replaced by an observer polynomial $T(q^{-1})$ chosen *a priori*. By replacing the polynomials in eqn. 2.1 with their estimates, the process model is given as follows ("(q^{-1})" is dropped from all polynomial representations for brevity),

$$\hat{A}y(t) = q^{-k}\hat{B}u(t) + d + T\xi(t) \tag{2.2}$$

where

$$\begin{aligned}
\hat{A} &= 1 + a_1q^{-1} + a_2q^{-2} + \cdots + a_{na}q^{na} \\
\hat{B} &= 1 + b_1q^{-1} + b_2q^{-2} + \cdots + b_{nb}q^{nb} \\
T &= 1 + t_1q^{-1} + t_2q^{-2} + \cdots + t_{nt}q^{nt}
\end{aligned}$$

Now consider the Diophantine identity:

$$T = E'_j\hat{A} + q^{-j}F'_j \tag{2.3}$$

which uniquely defines E'_j and F'_j when T, $\hat{A}$ and j are given. Multiplying eqn. 2.2 by E'_j and substituting eqn. 2.3 into eqn. 2.2 gives the following j^{th}-step ahead model,

$$Ty(t+j) = E'_j\hat{B}u(t+j-k) + F'_jy(t) + E'_jd + E'_jT\xi(t+j) \tag{2.4}$$

By applying the following identity to eqn. 2.4,

$$E'_j\hat{B} = G'_jT + H'_jq^{-j} \tag{2.5}$$

where the degrees of G'_j and H'_j are $j-1$ and $\max(nb-1, nt-1)$, past (known) control actions are separated from the unknown current and future values so that the optimal j^{th}-step ahead predictor is given as follows,

$$\hat{y}(t+j \mid t) = G'_j u(t+j-k) + H'_j u_f(t-k) + F'_j y_f(t) + r \tag{2.6}$$

where

$$\begin{aligned} u_f(t-k) &= \frac{u(t-k)}{T} \\ y_f(t) &= \frac{y(t)}{T} \\ r &= \frac{E'_j(1)}{T(1)} d \end{aligned}$$

The steady-state prediction term is obtained by applying the final value theorem to eqn. 2.2,

$$\hat{y}(s \mid t) = \frac{\hat{B}(1)}{\hat{A}(1)} u(t) + \frac{d}{\hat{A}(1)} \tag{2.7}$$

where s denotes a value at steady state. The expected values of the terms in $T\xi(\cdot)$ are zeros. Eqn. 2.7 represents a local linearization of the steady-state value of the process. This value is assumed to be valid in a local region around the current operating condition and will be continually updated to adapt to changing conditions.

2.2.2 Transfer Function Model with Integrator

When the process model assumes the form of an auto-regressive *integrated* moving-average (ARIMAX) with an exogenous input,

$$\hat{A} y(t) = q^{-k} \hat{B} u(t) + T \frac{\xi(t)}{\Delta} \tag{2.8}$$

where

$$\Delta = 1 - q^{-1}$$

a d.c. gain term is no long necessary as it is removed by the integrator Δ. Having the integrator in the noise model also represents a class of non-stationary dynamics such as random steps at random times, *i.e.* random-walk or Brownian motion type disturbances.

To derive a j^{th}-step ahead predictor from eqn. 2.8, another Diophantine identity is considered below,

$$T = E_j \hat{A} \Delta + q^{-j} F_j \tag{2.9}$$

Multiplying eqn. 2.8 by $\Delta E_j q^j$ gives:

$$E_j \hat{A} \Delta y(t+j) = E_j \hat{B} \Delta u(t+j-k) + E_j T \xi(t+j) \tag{2.10}$$

Again substituting eqn. 2.9 into eqn. 2.10 results in the following j^{th}-step ahead model,

$$T y(t+j) = E_j \hat{B} \Delta u(t+j-k) + F_j y(t) + E_j T \xi(t+j) \tag{2.11}$$

Future control actions are separated from the known ones by substituting the following identity into eqn. 2.11,

$$E_j \hat{B} = G_j T + H_j q^{-j} \tag{2.12}$$

where the elements of G_j now corresponds to the first $j-1$ open-loop step response coefficients. The optimal j^{th}-step ahead predictor becomes

$$\hat{y}(t+j \mid t) = G_j \Delta u(t+j-k) + H_j \Delta u_f(t-k) + F_j y_f(t) \tag{2.13}$$

where u_f and y_f are values filtered by $\frac{1}{T}$.

One might be tempted to calculate the steady-state output prediction by applying the final value theorem to the predictive model stated in eqn. 2.13. However, the presence of the integrator yields an indeterminate result. The following Lemma helps to explain why the steady-state prediction from an ARIMAX model is different from that of an ARMAX model.

Lemma 1

Given two polynomials T and A with degrees nt and na, for any integer value of j, polynomials E'_j and F'_j are uniquely defined by the following identity:

$$T = E'_j A + q^{-j} F'_j$$

Let r_i, $i = 1, 2, \cdots, \bar{na}$, denote the distinct roots of polynomial A so that multiplicity of r_i is s_i and $\sum_{i=1}^{\bar{na}} s_i = na$. If the roots of A are all inside the unit circle, i.e. $\mid r_i \mid < 1, \forall\, i = 1$ to na, polynomial E'_j converges to $\frac{T}{A}$ and F'_j to zero as j tends to infinity.

Proof

$\frac{1}{A}$ can be factored into the sum of $\bar{na}$ functions such that

$$\frac{1}{A} = f_1(q) + f_2(q) + \cdots + f_{\bar{na}}(q)$$

where

$$f_i(q) = \frac{\sigma_i}{(1 - r_i q^{-1})^{s_i}}$$

For each first-order factor, $\frac{1}{1-r_i q^{-1}}$, the Taylor series expansion around the point r_i has a remainder of zero in the limit as $n \to \infty$, *i.e.*

$$f_i(q) = \sigma_i \left(1 + r_i q^{-1} + (r_i q^{-1})^2 + \cdots + (r_i q^{-1})^n + \cdots\right)^{s_i}$$

and this series converges absolutely on the interval $\mid r_i \mid < 1$. Since $f_i(q)$ is a multiplication of s_i power series with the same interval of convergence, it also converges absolutely. According to the Taylor series expansion of functions, a convergent series can be added or subtracted to another convergent series to give yet another convergent series. Therefore, the Taylor series expansion of $\frac{T}{A}$ is given by E'_j and the remainder $F'_j q^{-j}$ approaches zero as j tends to infinity.

□□□

Remark 1

Recall the Diophantine identity for the ARIMAX model in eqn. 2.9,

$$T = E_j \bar{A} + q^{-j} F_j \tag{2.14}$$

where

$$\bar{A} = A\Delta$$

The divisor polynomial, $\bar{A}$, has one root on the unit circle. Therefore, the remainder term $q^{-j}F_j$ does not vanish and E_j does **not** *approach $\frac{T}{A\Delta}$ as j increases to infinity.*

□□□

Remark 2

The term $\frac{T}{A\Delta}$ in eqn. 2.14 is similar to having a step input to the transfer function $\frac{T}{A}$ because $\frac{1}{\Delta}$ is a step function. The coefficients in the non-convergent polynomial E_j are the step response model parameters that, for sufficiently large j, approach the steady-state gain, $\frac{T(1)}{A(1)}$. Since the series E_j converges to an infinite stationary series (assuming all roots of A are inside the unit circle), the remainder polynomial, F_j which is of finite order, necessarily converges to a finite constant polynomial (defined as F_s).

□□□

Remark 3

The polynomials G_j and H_j in eqn. 2.12 have a similar convergence to E_j and F_j. Substituting eqn. 2.12 into eqn. 2.9 gives

$$\frac{B}{A\Delta} = G_j + q^{-j}\left(\frac{H_j}{T} + \frac{BF_j}{A\Delta T}\right) \tag{2.15}$$

Since F_j converges to F_s and G_j corresponds to an infinite series of step response model parameters in the limit as j approaches infinity, H_j has to converge to a constant polynomial also. Therefore, for sufficiently large j,

$$q^j\left(\frac{B}{A\Delta} - G_j\right) = \frac{H_s}{T} + \frac{BF_s}{A\Delta T} \tag{2.16}$$

□□□

Remark 4

The term $\frac{B}{A\Delta}$ in eqn. 2.16 is also similar to having a step input to the transfer function $\frac{B}{A}$. The elements in polynomial G_j correspond to the step response coefficients. Therefore, the left-hand-side of eqn. 2.16, for sufficiently large j, can be further reduced to $\frac{g_s}{\Delta}$ such that

$$\frac{H_s}{T} + \frac{BF_s}{A\Delta T} = \frac{g_s}{\Delta} \tag{2.17}$$

□□□

Now applying the result in Lemma 1 to the j^{th}-step ARMAX model in eqn. 2.4,

$$\lim_{j\to\infty} Ty(t+j) = \frac{T}{\hat{A}}\hat{B}u(t+j-k) + \frac{T}{\hat{A}}d + \frac{T^2}{\hat{A}}\xi(t+j) \tag{2.18}$$

the optimal steady-state prediction at time t is given as

$$\lim_{j\to\infty} \hat{y}(t+j) = \hat{y}(s \mid t) = \frac{\hat{B}(1)}{\hat{A}(1)}u(t) + \frac{d}{\hat{A}(1)} \tag{2.19}$$

which yields the same result as eqn. 2.7.

However, in the ARIMAX model, the remainder term F_j in eqn. 2.9 does not diminish as j increases. The future prediction of $\hat{y}(t+j)$ is always dependent on the term $F_j\ y_f(t)$ in which the coefficients of F_j change with time. Therefore, the output steady-state prediction $\hat{y}(s)$ requires the converged polynomial F_s. One could calculate the asymptotic steady-state prediction at the cost of a relatively heavy computation by iterating the Diophantine identities in eqns. 2.9 and 2.12 continually until $\hat{y}(t+j)$ in eqn. 2.13 becomes stationary. But an alternative approach at a much lesser cost is available by using the results in Remarks 2 and 3. It allows the open-loop steady-state value to be calculated by directly computing the polynomials F_s and H_s.

Now recall that polynomials F_{j+1} and H_{j+1} are defined as follows:

$$T = E_{j+1}\hat{A}\Delta + q^{-j-1}F_{j+1} \tag{2.20}$$

$$E_{j+1}\hat{B} = G_{j+1}T + q^{-j-1}H_{j+1} \tag{2.21}$$

Subtracting eqns. 2.9 and 2.12 from eqns. 2.20 and 2.21 respectively gives

$$e_j\hat{A}\Delta = F_j - q^{-1}F_{j+1} \tag{2.22}$$

$$e_j\hat{B} = g_jT + q^{-1}H_{j+1} - H_j \tag{2.23}$$

where F and H are of degree na and $\max(nb, nt) - 1$ respectively, and e_j and g_j are the last coefficients in E_{j+1} and G_{j+1}. For sufficiently large j,

$$e_j\hat{A}\Delta = e_s\hat{A}\Delta = F_s - q^{-1}F_s \tag{2.24}$$

$$e_j\hat{B} = e_s\hat{B} = g_sT + q^{-1}H_s - H_s \tag{2.25}$$

where (from Remark 2)

$$e_s = \frac{T(1)}{\hat{A}(1)} \tag{2.26}$$

$$g_s = \frac{\hat{B}(1)}{\hat{A}(1)} \tag{2.27}$$

Therefore,

$$F_s = e_s\hat{A} \tag{2.28}$$

$$H_s\Delta = g_sT - e_s\hat{B} \tag{2.29}$$

$$H_s = h_{s,0} + h_{s,1}q^{-1} + \cdots + h_{s,nh}q^{-nh} \tag{2.30}$$

where

$$\begin{aligned} nh &= \max(nb-1, nt-1) \\ h_{s,i} &= e_s \sum_{j=i+1}^{nb} b_j - g_s \sum_{j=i+1}^{nt} t_j \end{aligned}$$

The optimal steady-state predictor based on an ARIMAX model is then given as

$$\hat{y}(s|t) = g_s \sum_{j=1}^{nu} \Delta u(t+j-k) + H_s \Delta u_f(t-k) + F_s y_f(t) \tag{2.31}$$

where nu is the control horizon.

The first term in the right hand side of eqn. 2.31 is the current and future forced response whereas the remaining terms correspond to the impact of past inputs on the steady-state value of the process output.

2.2.3 Convolution Model

The convolution model structure is usually represented by a step response model:

$$y(t) = \sum_{i=1}^{\infty} g_i q^{-i} \Delta u(t+1-k) + a_{ss} \tag{2.32}$$

where the step response coefficients, g_i, tend to g_s when i tends to infinity and a_{ss} is a reference value. It is readily turned into an impulse response model by differencing the step response such that

$$h_i = \Delta g_i$$

In practice, the dynamic model is commonly assumed by a finite number of coefficients. *i.e.*

$$y(t) = \sum_{i=1}^{T} g_i q^{-i} \Delta u(t+1-k) + a_e + a_{ss} \tag{2.33}$$

where

T is the model horizon

$$a_e = \sum_{i=T}^{\infty} g_i q^{-i} \Delta u(t+1-k)$$

T is usually chosen so that the truncation error is less than 1 % of the step-input steady-state gain. Depending on the sampling rate, the number of coefficients may be as many as 50 to encompass 99 % of the step response dynamics. The process model becomes

$$y(t) = \sum_{i=1}^{T} g_i q^{-i} \Delta u(t+1-k) + a_{ss} \qquad (2.34)$$

Its advantage over a parametric model is the capability to model unusual dynamic behavior which cannot be well represented by a reduced-order model. Moreover, many chemical processes are inherently infinite order and therefore a simple first or second order transfer function may introduce structural mismatch that is unfavorable to predictive control (Shook *et al.*, 1990). However, on-line identification of such a large number of parameters is a difficult task.

The process model in eqn. 2.34 can be arranged into different predictive representations such as recursive form or state-space formulation (Lim, 1988; Li *et al.*, 1989; Garcia and Morari, 1982). But all of them assume a linear superposition of previous inputs to the total output such that (for unity time delay *i.e.* $k = 1$)

$$\hat{y}(t+j|t) = \sum_{i=1}^{j} g_i \Delta u(t+j-i) + \sum_{i=j+1}^{T} g_i \Delta u(t+j-i) + a_{ss} \qquad (2.35)$$

The output comprises three terms: future control effects, past control effects, and a reference term. Eqn. 2.35 shows that even with steady-state weighting a complete step response sequence of up to T intervals is required regardless of the choice of output prediction, j. However, an approximation alternative, borrowed from Auslander *et al.*(1978), can be used so that a finite number of initial step response coefficients plus the steady-state gain are sufficient for long-range predictions. This approximation recognizes that most process responses are dominated by a single time constant at least in the low frequency region, *i.e.* for overdamped systems, the low frequency dynamics can be well approximated by a first-order model.

Consider that a unit step response from a stable second or higher order process is often characterized by an S-shaped curve, the low frequency portion can be approximated by an exponential function from the n^{th} point to the steady state. In terms of impulse response, the approximation is written as

$$\begin{aligned} G_p &= h_1 q^{-1} + h_2 q^{-2} + \cdots + h_{n-1} q^{-n+1} + h_n q^{-n} + \cdots \\ &= h_1 q^{-1} + h_2 q^{-2} + \cdots + h_{n-1} q^{-n+1} + \frac{h_n q^{-n}}{1 - pq^{-1}} \end{aligned} \qquad (2.36)$$

where p, being the rate of exponential decay, is a discrete characterization of the dominant time constant. If the steady-state gain of the process is known, for example, from steady-state design equations (Chesna and Ydstie, 1986), p is determined so that the model gain in eqn. 2.36 is equal to the process gain.

$$\begin{aligned} G_p(1) &= g_s \\ p &= 1 - \frac{h_n}{g_s - \sum_{i=1}^{n-1} h_i} \end{aligned} \qquad (2.37)$$

The model in eqn. 2.36 can be re-arranged into a transfer function form

$$G_p = \frac{b'_1 q^{-1} + b'_2 q^{-2} + \cdots + b'_n q^{-n}}{1 - pq^{-1}} \qquad (2.38)$$

where

$$\begin{aligned} b'_1 &= h_1 \\ b'_i &= h_i - ph_{i-1} \quad \text{for } i = 1 \ \ to \ \ n \end{aligned}$$

A stable overdamped model is guaranteed as long as

$$g_s > \sum_{i=1}^{n} h_i \qquad (2.39)$$

It was shown that the modeling error stemmed from the exponential decay approximation can be made arbitrarily small by increasing the sampling period for a fixed n coefficients or vice versa (Takahashi *et al.*, 1975).

By transforming a "limited" convolution model (a few step response coefficients and gain) into a transfer function form, the ARIMAX modeling procedure in Section 2.2.2 can be applied to make future predictions. In so doing, the user-defined polynomial T is selected to be $\hat{A} = (1 - pq^{-1})$ such that the ARIMAX modeling conforms to an output error prediction type which has been the case for algorithms such as DMC and MOCCA using convolution model type.

2.3 Control Algorithm

2.3.1 Derivation

The design of a long-range, model-predictive controller is based on the process predicted behavior over the prediction horizon. The control objective is to have the predicted values track a setpoint trajectory within the horizon. The following quadratic function defines the performance index to be minimized to achieve the control objective.

$$\begin{aligned} J(n1, n2, nu, \lambda, \gamma) &= \sum_{j=n1}^{n2} \gamma_y(j) \left[\hat{y}(t+j) - w(t+j)\right]^2 + \sum_{j=1}^{nu} \lambda(j) \left[\Delta u(t+j-1)\right]^2 \\ &\quad + \sum_{j=1}^{nu} \gamma(j) \left[\hat{y}(s|t+j-1) - w(s)\right]^2 \end{aligned} \tag{2.40}$$

where

- $n1$ is the minimum output prediction horizon,
- $n2$ is the maximum output prediction horizon,
- nu is the control horizon such that $\Delta u(t+j) = 0, \forall\, j \geq nu$,
- $\gamma_y(j)$ is an output weighting sequence,
- $\lambda(j)$ is a control weighting sequence,

- $\gamma(j)$ is a steady-state error weighting sequence,

- s denotes a value at steady state.

The first two terms on the right-hand side form the standard Generalized Predictive Control (GPC) objective. The last term corresponds to the additional terms penalizing the squares of errors at the predicted steady state. The summation from 1 to nu is required because, at each control interval j where $j = 1$ to nu, the steady-state prediction involves the sum of first j consecutive control actions. It has been suggested that, for robust control, $n2$ should encompass the whole rise time of a process which often requires as many as 50 terms. Now with steady-state error weighting in place, $n2$ can be reduced to cover the initial high frequency dynamics.

The process predicted values are required in eqn. 2.40 from $j = n1$ to $n2$ for $\hat{y}(t+j)$ and $j = 1$ to nu for $\hat{y}(s|t+j-1)$. To derive a control law from eqn. 2.40, the predictive equations for the ARIMAX model in eqns. 2.13 and 2.31 are used. The ARIMAX model can be put into the following compact vector/matrix equation (assuming unity delay):

$$\mathbf{Y} = \mathbf{G}\mathbf{U} + \mathbf{f} \tag{2.41}$$

where

$$\begin{aligned}
\mathbf{Y} &= \begin{bmatrix} \hat{y}(t+n1) & \hat{y}(t+n1+1) & \cdots & \hat{y}(t+n2) \end{bmatrix}^T && (2.42) \\
\mathbf{G} &= \begin{bmatrix}
g_{n1-1} & \cdots & g_0 & 0 & 0 & \cdots & 0 \\
g_{n1} & \cdots & g_1 & g_0 & 0 & \cdots & 0 \\
\vdots & \ddots & & & \ddots & & \vdots \\
\vdots & & \ddots & & & & g_0 \\
\vdots & \cdots & & & & & \vdots \\
g_{n2-1} & g_{n2-2} & \cdots & & & & g_{n2-nu}
\end{bmatrix} && (2.43) \\
\mathbf{U} &= \begin{bmatrix} \Delta u(t+n1-1) & \Delta u(t+n1) & \cdots & \Delta u(t+n2-1) \end{bmatrix}^T && (2.44)
\end{aligned}$$

$$\mathbf{f} = \left[\; f(t+n1) \quad f(t+n1+1) \quad \cdots \quad f(t+n2) \;\right]^T \tag{2.45}$$

$$f(t+j) = H_j\, \Delta u_f(t-1) + F_j\, y_f(t) \tag{2.46}$$

and

$$\mathbf{Y_s} = \mathbf{G_s U} + \mathbf{f_s} \tag{2.47}$$

where

$$\mathbf{Y_s} = \left[\; \hat{y}(s|t) \quad \hat{y}(s|t+1) \quad \cdots \quad \hat{y}(s|t+nu-1) \;\right]^T \tag{2.48}$$

$$\mathbf{G_s}\ (nu \text{ x } nu) = \begin{bmatrix} g_s & 0 & \cdots & 0 \\ g_s & g_s & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ g_s & \cdots & \cdots & g_s \end{bmatrix} \tag{2.49}$$

$$\mathbf{f_s}\ (nu \text{ x } 1) = \left[\; 1 \quad 1 \quad \cdots \quad 1 \;\right]^T \cdot \left(H_s\, \Delta u_f(t-k) + F_s\, y_f(t)\right) \tag{2.50}$$

The equivalent of eqn. 2.40 using vector/matrix notation is:

$$J = [\mathbf{Y} - \mathbf{W}]^T \Gamma_y [\mathbf{Y} - \mathbf{W}] + \mathbf{U}^T \Lambda \mathbf{U} + [\mathbf{Y_s} - \mathbf{W_s}]^T \Gamma [\mathbf{Y_s} - \mathbf{W_s}] \tag{2.51}$$

where

$$\begin{aligned} \Gamma_y &= \operatorname{diag}\left[\gamma_y(n1) \;\; \gamma_y(n1+1) \;\; \cdots \;\; \gamma_y(n2)\right] \\ \Lambda &= \operatorname{diag}\left[\lambda(1) \;\; \lambda(2) \;\; \cdots \;\; \lambda(nu)\right] \\ \Gamma &= \operatorname{diag}\left[\gamma(1) \;\; \gamma(2) \;\; \cdots \;\; \gamma(nu)\right] \\ \mathbf{W} &= \left[w(t+n1) \;\; w(t+n1+1) \;\; \cdots \;\; w(t+n2)\right]^T \\ \mathbf{W_s}\ (nu \text{ x } 1) &= \left[1 \;\; 1 \;\; \cdots \;\; 1\right]^T \cdot w(s) \end{aligned}$$

After substituting eqns. 2.41 and 2.47 into eqn. 2.51, the cost function is minimized by the following control equation:

$$\mathbf{U} = \left[\mathbf{G}^T \Gamma_y \mathbf{G} + \Lambda + \mathbf{G_s}^T \Gamma \mathbf{G_s}\right]^{-1} \left[\mathbf{G}^T \Gamma_y (\mathbf{W} - \mathbf{f}) + \mathbf{G_s}^T \Gamma (\mathbf{W_s} - \mathbf{f_s})\right] \tag{2.52}$$

where $\mathbf{G}^T\Gamma_y\mathbf{G}$ and ${\mathbf{G_s}}^T\Gamma\mathbf{G_s}$ are both matrices of dimension nu x nu. The dynamic matrix, $\mathbf{G}$, contains all the step response coefficients arranged in a lower triangular structure. The term ${\mathbf{G_s}}^T\Gamma\mathbf{G_s}$ is always of full rank. It guarantees a non-singular inverse even if the matrix $\mathbf{G}^T\Gamma_y\mathbf{G}$ is ill-conditioned due to large time delays or to too short an output prediction horizon. When Γ is set to zero, eqn. 2.52 reduces to the basic GPC control law. Therefore, eqn. 2.52 is referred in short as "GPC with γ-weighting". For simplicity, the output weighting is assumed to be unity; thus Γ_y reduces to an identity matrix in the sequel.

The control law in eqn. 2.52 is derived specifically with integral control action, *i.e.* the ARIMAX model. However, similar control laws can be obtained for the other two modeling methods described in Sections 2.2.1 and 2.2.3.

2.3.2 Control Law in Linear Form

Although the control law in eqn. 2.52 calculates the future control action from $t = 1$ to nu, only the current control action, i.e. $\Delta u(t)$, is implemented. At the next sample, the entire long-range minimization is repeated once again to calculate the next control action. This receding-horizon policy produces better control performance as the control makes use of updated plant information. It has been shown that, in the absence of disturbance or model change, the control sequence in successive times are the same (Clarke and Mohtadi, 1989). Because of the receding-horizon policy, only the first element in vector $\mathbf{U}$ of eqn. 2.52 is required. Therefore, writing the control law in a linear difference equation form enables the analysis of this controller in the closed-loop form (Lambert, 1987; McIntosh *et al.*, 1991).

A general linear structure of a controller with a setpoint and measurement variable is represented by

$$R\Delta u(t) = Vw(t) - Sy(t) \tag{2.53}$$

where R, V, and S are polynomials in the backward shift operator.

Now define a vector

$$\left[\ \alpha_{n1} \quad \alpha_{n1+1} \quad \cdots \quad \alpha_{n2}\ \right] \tag{2.54}$$

which denotes the first row of the matrix $[\mathbf{G}^T\mathbf{G} + \Lambda + \mathbf{G_s}^T\Gamma\mathbf{G_s}]^{-1}\mathbf{G}^T$, and a variable

$$\alpha_s = \beta_{n1} + \beta_{n1+1} + \cdots + \beta_{n2} \tag{2.55}$$

where β_i corresponds to the elements of the first row vector in the matrix

$$\left[\mathbf{G}^T\mathbf{G} + \Lambda + \mathbf{G_s}^T\Gamma\mathbf{G_s}\right]^{-1}\mathbf{G_s}^T\Gamma$$

Then the current control action can be written as follows, (assuming all future set-points are the same as the present one)

$$\Delta u(t) = \tag{2.56}$$

$$[\alpha_{n1}\ \alpha_{n1+1}\ \cdots\ \alpha_{n2}\ \alpha_s] \begin{bmatrix} w(t) & -F_{n1}\, y_f(t) & -H_{n1}\, \Delta u_f(t-1) \\ w(t) & -F_{n1+1}\, y_f(t) & -H_{n1+1}\, \Delta u_f(t-1) \\ \vdots & \vdots & \vdots \\ w(t) & -F_{n2}\, y_f(t) & -H_{n2}\, \Delta u_f(t-1) \\ w(t) & -F_s\, y_f(t) & -H_s\, \Delta u_f(t-1) \end{bmatrix}$$

and in compact form,

$$\begin{aligned} \Delta u(t) \;&=\; w(t)\sum_{i=n1}^{n2}\alpha_i - y_f(t)\sum_{i=n1}^{n2}F_i\alpha_i - \Delta u_f(t-1)\sum_{i=n1}^{n2}H_i\alpha_i \\ &\quad + w(t)\,\alpha_s\ - y_f(t)F_s\alpha_s - \Delta u_f(t-1)H_s\alpha_s \end{aligned} \tag{2.57}$$

By comparing the terms in eqn. 2.57 with eqn. 2.53, the polynomials in eqn. 2.53 can be shown to be as follows,

$$\begin{aligned} R &= T + q^{-1}\left(\sum_{i=n1}^{n2} H_i\alpha_i + H_s\alpha_s\right) & (2.58)\\ V &= T\left(\sum_{i=n1}^{n2} \alpha_i + \alpha_s\right) & (2.59)\\ S &= \sum_{i=n1}^{n2} F_i\alpha_i + F_s\alpha_s & (2.60)\end{aligned}$$

where the degrees are $\delta R = \max(nb, nt)$, $\delta V = nt$, and $\delta S = na$.

2.4 Closed-Loop Analysis of γ-Weighting

2.4.1 Closed-Loop Transfer Function

A closed-loop system with a general linear controller represented by eqns. 2.58 to 2.60 combined with the ARIMAX model in eqn. 2.8 is depicted in fig. 2.2. The non-stationary noise signal $x(t)$ is represented by

$$x(t) = \frac{\xi(t)}{\Delta} \tag{2.61}$$

where $\xi(t)$ is a white noise sequence and Δ is the difference operator.

Assuming T to be 1, the designed closed-loop transfer functions for $y(t)$ and $u(t)$ are given by

$$y(t) = \frac{\hat{B}Vq^{-1}w(t) + R\Delta x(t)}{R\hat{A}\Delta + q^{-1}\hat{B}S} \tag{2.62}$$

$$u(t) = \frac{\hat{A}Vw(t) - Sx(t)}{R\hat{A}\Delta + q^{-1}\hat{B}S} \tag{2.63}$$

When there is a model-plant mismatch, particularly if the gain of the model is different from the actual gain of the plant due to inexact estimation of polynomials A and B, the presence of the steady-state error weighting term does not cause any

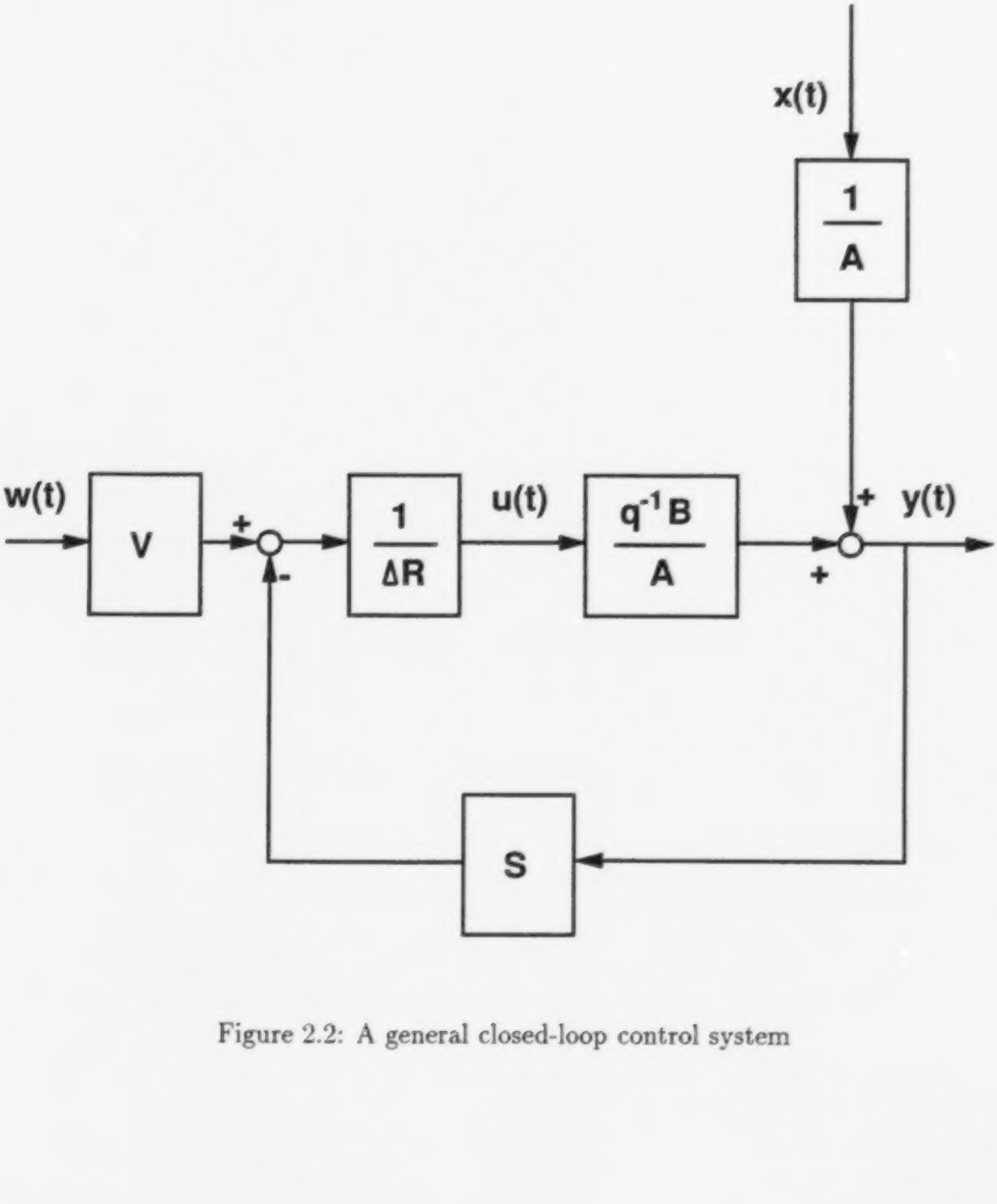
x(t)
1/A
w(t)
V
+
-
1/ΔR
u(t)
q^-1 B / A
+
+
y(t)
S

Figure 2.2: A general closed-loop control system

offset in the closed-loop system, *i.e.* it ensures that the sign of the gain is correct. This property can be seen by analyzing the closed-loop transfer function at steady state.

At steady state (i.e. $q = 1$), the closed-loop transfer function in eqn. 2.62 is reduced to

$$y(t) = \frac{V(1)}{S(1)} w(t) \tag{2.64}$$

where

$$\begin{aligned} V(1) &= \sum_{i=n1}^{n2} \alpha_i + \alpha_s \\ S(1) &= \sum_{i=n1}^{n2} \alpha_i F_i(1) + \alpha_s F_s(1) \end{aligned} \tag{2.65}$$

Recall from eqn. 2.9 for $T = 1$ that,

$$1 = E_j \hat{A} \Delta + q^{-j} F_j \tag{2.66}$$

At steady state, it becomes

$$1 = F_j(1) = F_s(1) \tag{2.67}$$

Therefore,

$$V(1) = S(1) = \sum_{i=n1}^{n2} \alpha_i + \alpha_s \tag{2.68}$$

confirms that

$$y(t) = w(t) \tag{2.69}$$

at steady state, and for step-type inputs,

$$\frac{y(t)}{x(t)} = 0$$

2.4.2 Minimum Variance Control with γ-Weighting

For the special case when the following settings are used:

$$\left.\begin{array}{rcl} n1 &=& 1 \\ n2 &=& k \\ nu &=& 1 \\ \lambda &=& 0 \\ \gamma &=& 0 \end{array}\right\} \tag{2.70}$$

the control law from eqn. 2.57 can be written as

$$\sum_{i=1}^{k} g_{i-1}^2 \Delta u(t) = \sum_{i=1}^{k} g_{i-1} \left[w(t) - F_i y(t) - H_i \Delta u(t-1) \right] \tag{2.71}$$

where g_i's are the step response coefficients. Since the first $k-1$ step response coefficients are zero due to the k^{th}-step time delay, eqn. 2.71 is reduced to

$$\Delta u(t) = \frac{w(t) - F_k\, y(t)}{g_{k-1} + q^{-1} H_k} \tag{2.72}$$

Substituting eqn. 2.12 for $j = k$ and $T = 1$ into eqn. 2.72 turns the control law into the following familiar form

$$\Delta u(t) = \frac{w(t) - F_k\, y(t)}{\hat{B}'\, E_k} \tag{2.73}$$

where

$$q^{-k+1} \hat{B}' = \hat{B} \tag{2.74}$$

which, for $w(t) = 0$, is the well-known minimum-variance regulator (Åström and Wittenmark, 1973).

From the equation above, the closed-loop characteristic polynomial is obtained as

$$\begin{aligned} \mathcal{C}_{MV} &= [g_{k-1} + q^{-1} H_k]\, \hat{A}\Delta + q^{-1} \hat{B}\, F_k \\ &= \Delta\hat{A}\, g_{k-1} + \Delta\hat{A}\, q^{-1} \left[H_k + \frac{\hat{B}}{\Delta\hat{A}} F_k \right] \end{aligned} \tag{2.75}$$

Substituting eqn. 2.15 for $j = k$ and $T = 1$ into eqn. 2.75 gives

$$\begin{aligned}\mathcal{C}_{MV} &= \Delta\hat{A}\, g_{k-1} + \Delta\hat{A}\, \left[\frac{\hat{B}}{\Delta\hat{A}} - g_{k-1}\right] q^{k-1} \\ &= \hat{B}'\end{aligned} \tag{2.76}$$

This confirms that the basic GPC algorithm with the settings in eqn. A.5 is reduced to the minimum variance control which assigns all closed-loop poles to the positions of open-loop zeros. In practice, this controller is not usually acceptable in process control application because of large and vigorous control actions as well as possible cancellation of open-loop unstable zeros in the closed-loop transfer function.

For the special case with steady-state error weighting, *i.e.* with the following settings:

$$\left.\begin{array}{rcl} n1 &=& 1 \\ n2 &=& k \\ nu &=& 1 \\ \lambda &=& 0 \\ 0 < \gamma &<& \infty \end{array}\right\} \tag{2.77}$$

the closed-loop characteristic polynomial becomes

$$\mathcal{C}_{MV,S} = R\,\hat{A}\,\Delta + q^{-1}\,\hat{B}\,S \tag{2.78}$$

where the polynomials other than the models are given from eqns. 2.58 and 2.60 as

$$R = 1 + q^{-1}\frac{g_{k-1}H_k + g_s H_s \gamma}{(g_{k-1}^2 + g_s^2\gamma)} \tag{2.79}$$

$$S = 1 + \frac{g_{k-1}F_k + g_s F_s \gamma}{(g_{k-1}^2 + g_s^2\gamma)} \tag{2.80}$$

After grouping the steady state terms together, eqn. 2.78 becomes

$$\begin{aligned}\mathcal{C}_{MV,S} &= g_{k-1}\left[\hat{A}\Delta(g_{k-1} + q^{-1}H_k\,) + q^{-1}\hat{B}F_k\right] \\ &+ g_s\gamma\left[\hat{A}\Delta(g_s + q^{-1}H_s\,) + q^{-1}\hat{B}F_s\right]\end{aligned} \tag{2.81}$$

Substituting eqn. 2.15 for $j = k$ and $T = 1$ and eqn. 2.17 into eqn. 2.81, the characteristic polynomial in its simplified form emerges as

$$C_{MV,S} = \hat{B}' + \frac{g_s^2}{g_{k-1}} \gamma \hat{A} \tag{2.82}$$

Stability is determined by its roots which can be examined by using root-locus analysis for various γ values. Eqn. 2.82 is similar in structure to the characteristic equation of an ordinary control weighted one-step ahead controller, implying that the role of γ-weighting on the steady-state error is related to the λ-weighting on control actions. To find the relationship between the two, one can examine the γ-weighting in eqn. 2.40.

$$J(\gamma) = \gamma \sum_{j=1}^{nu} [\hat{y}(s|t+j-1) - w(s)]^2 \tag{2.83}$$

Substituting eqn. 2.31 into eqn. 2.83 and expanding the quadratic term gives:

$$\begin{aligned} J(\gamma) \;\; = \;\; & \gamma g_s^2 \sum_{j=1}^{nu} \left(\sum_{i=1}^{j} \Delta u(t+i-1) \right)^2 \\ & + 2 g_s \gamma \left(f(s) - w(s) \right) \sum_{j=1}^{nu} \left(\sum_{i=1}^{j} \Delta u(t+i-1) \right) \\ & + \gamma \left(f(s) - w(s) \right)^2 \end{aligned} \tag{2.84}$$

which shows that γ penalizes not only the squares of incremental control action as λ does, but also the linear product of the total forced output $g_s^2 \sum_{j=1}^{nu} \sum_{i=1}^{j} \Delta u(t+i-1)$ and the open-loop free response $(f(s) - w(s))$. The best possible minimum of $J(\gamma)$ is to have the two terms opposite in sign. Therefore, the linear product term can be interpreted as a weighting of $2\gamma g_s(f(s) - w(s))$ in the direction of Δu such that a positive gain process with a projected steady-state value larger than the setpoint would favor a negative control increment to bring down the steady-state value. Thus, the use of steady-state error weighting provides a gain-dependent control weighting compared to ordinary control weighting which indiscriminately penalizes control moves. Its advantage can also be substantiated by considering the comparisons to follow.

Three closed-loop transfer functions using three different modeling schemes and control laws for the same process are compared.

Case 1 ARIMAX model with weighted one-step ahead control

model:

$$Ay(t) = Bu(t-k) + \frac{x(t)}{\Delta}$$

control law:

$$J_1 = [y(t+k) - w(t+k)]^2 + [g_s \Delta u(t)]^2$$

The resulting closed-loop transfer function is

$$y(t) = \frac{g_{k-1} B q^{-1} w(t) + [g_{k-1} E_k B + g_s^2] x(t)}{g_{k-1} B + g_s^2 \Delta A}$$

Although it provides offset-free servo control, an extra order in the denominator is incurred by the integrated mode causing a larger phase shift in frequency domain which requires a lower gain and hence more sluggish response.

Case 2 ARMAX model with weighted one-step ahead control

model:

$$Ay(t) = Bu(t-k) + x(t)$$

control law:

$$J_2 = [y(t+k) - w(t+k)]^2 + [g_s u(t)]^2$$

In the absence of the integrated mode, the modeling of the noise is assumed to be always stationary. The closed-loop transfer function is

$$y(t) = \frac{g_{k-1}Bq^{-1}w(t) + [g_{k-1}E_kB + g_s^2]x(t)}{g_{k-1}B + g_s^2 A}$$

The characteristic polynomial does not have an extra order. However, this approach is well known for having offset problem at steady state.

Case 3 ARIMAX model with one-step ahead control and steady-state error weighting

model:

$$Ay(t) = Bu(t-k) + \frac{x(t)}{\Delta}$$

control law:

$$J_3 = [y(t+k) - w(t+k)]^2 + \gamma\,[y_{s,t} - w_s]^2$$

transfer function:

$$y(t) = \frac{(g_{k-1} + \gamma g_s)q^{-1}Bw(t) + [g_{k-1}E_kB + \gamma g_s q^{-1}H_s + g_s^2]x(t)}{g_{k-1}B + \gamma g_s^2 A}$$

The introduction of weighting on the steady-state error instead of on the control increments eliminates both the extra order in Case 1 and the offset problem in Case 2. (See Appendix A for a step-by-step derivation.)

The reason why the characteristic equation for Case 3 is lower in order is due to the weighting on the linear product term in eqn. 2.84. The minimization of eqn. 2.84 with respect to Δu gives a γ-weighted linear Δu term and a γ-weighted linear $(f(s) - w(s))$. Their combination in the closed-loop characteristic equation results in γA which is free of the integrator. Without the linear product term, eqn. 2.84 would be equivalent to an ordinary control weighting with $\lambda\Delta A$ in the closed-loop.

2.4.3 Mean-Level Property

For open-loop stable process, the controller with the same settings as in eqn. A.5 except for $n2 \rightarrow \infty$ tends to be a mean-level controller, which has the control law of the following form,

$$\Delta u(t) = g_s^{-1}[w(t) - f(t+\infty)] \tag{2.85}$$

where

- g_s represents the process gain,
- $f(t+\infty)$ is the open-loop process free response at infinite time.

Obtaining this type of controller is not practical in digital control because it requires a large number of iterations of the diophantine identity or unforced output predictions. Analogous to $n2 \rightarrow \infty$ is $\gamma \rightarrow \infty$ in eqn. 2.40. Therefore, a large weighting on the steady state error regardless of the size of output prediction horizon should be able to approximate the mean-level control.

It is known that mean-level control assigns all closed-loop poles to the same positions as the open-loop poles so that a single step change in control due to a step in the setpoint will take the output to the setpoint as in an open-loop step change. Now consider the elements in vector 2.54. For the case of no weighting on the control increments, the elements are

$$\alpha_i = \frac{g_{i-1}}{\sum_{j=n1}^{n2} g_{j-1}^2 + \gamma g_s^2} \qquad \text{for } i = n1 \textit{ to } n2 \tag{2.86}$$

$$\alpha_s = \frac{\gamma g_s}{\sum_{j=n1}^{n2} g_{j-1}^2 + \gamma g_s^2} \tag{2.87}$$

As the value of γ increases, α_i diminishes and α_s approaches $\frac{1}{g_s}$. Therefore, substituting eqns. 2.86 and 2.87 into eqns. 2.58 to 2.59 for sufficiently large γ yields

$$\begin{aligned} R &= 1 + q^{-1}\left(\sum_{i=n1}^{n2} H_i\alpha_i + H_s\alpha_s\right) \\ &\approx 1 + q^{-1}\frac{H_s}{g_s} \end{aligned} \tag{2.88}$$

$$\begin{aligned} V &= \sum_{i=n1}^{n2} \alpha_i + \alpha_s \\ &\approx \frac{1}{g_s} \end{aligned} \tag{2.89}$$

$$\begin{aligned} S &= \sum_{i=n1}^{n2} F_i\alpha_i + F_s\alpha_s \\ &\approx \frac{F_s}{g_s} \end{aligned} \tag{2.90}$$

The closed-loop characteristic equation becomes

$$\mathcal{C}_s = [R\hat{A}\Delta + q^{-1}\hat{B}S] \tag{2.91}$$

$$\begin{aligned} &= [g_s + q^{-1}H_s]\hat{A}\Delta + q^{-1}\hat{B}F_s \\ &= \hat{A}\left[\Delta g_s + \Delta q^{-1}\left(H_s + \frac{\hat{B}F_s}{\hat{A}\Delta}\right)\right] \end{aligned} \tag{2.92}$$

Substituting eqn. 2.17 with $T = 1$ into eqn. 2.92 for $j \to \infty$ gives

$$\mathcal{C}_s = \hat{A}\, g_s \tag{2.93}$$

which confirms that $\gamma \to \infty$ assigns all closed-loop poles to the positions of those for open-loop. This stabilizing ability of γ is formally stated in the following lemma.

Lemma 2

For any open-loop stable process, $G_p(q^{-1})$, under long-range predictive control with a steady-state weighting term, γ, and a stable model $\hat{G}_p(q^{-1})$, there is a finite value, γ_m, such that $\gamma > \gamma_m > 0$ guarantees a stable closed-loop system.

Proof

The closed-loop characteristic equation from eqns. 2.58 and 2.60 and eqn. 2.91 without taking the limit is given as:

$$\begin{aligned} \mathcal{C}_s &= \left[1 + q^{-1}\left(\sum_{j=n1}^{n2} H_j\alpha_j + H_s\alpha_s\right)\right]\hat{A}\Delta + q^{-1}\hat{B}\left(\sum_{j=n1}^{n2} F_j\alpha_j + F_s\alpha_s\right) \\ &= \hat{A}\Delta + \hat{A}\Delta q^{-1}\left[\sum_{j=n1}^{n2}\left(H_j + \frac{\hat{B}F_j}{\hat{A}\Delta}\right)\alpha_j + \left(H_s + \frac{\hat{B}F_s}{\hat{A}\Delta}\right)\alpha_s\right] \qquad (2.94) \end{aligned}$$

By using the results in Remarks 2 to 4, eqn. 2.94 becomes

$$\mathcal{C}_s = \hat{A}\Delta + \hat{A}\Delta q^{-1}\left[\sum_{j=n1}^{n2}\left(\frac{\hat{B}}{\hat{A}\Delta} - G_j\right)q^j\alpha_j + \left(\frac{g_s}{\Delta}\right)\alpha_s\right] \qquad (2.95)$$

Substituting eqns. 2.86 and 2.87 into eqn. 2.95 yields

$$\mathcal{C}_s = \hat{A} + \mathcal{D} \qquad (2.96)$$

where

$$\begin{aligned} \mathcal{D} &= \frac{\hat{A}q^{-1}}{\sum_{i=n1}^{n2} g_{i-1}^2 + \gamma g_s^2}\left[\Delta\sum_{j=n1}^{n2}\left(g_j + g_{j+1}q^{-1} + \cdots\right)g_{j-1} + g_s^2\gamma - 1\right] \qquad (2.97) \\ &= \delta_1 q^{-1} + \delta_2 q^{-2} + \cdots \end{aligned}$$

The coefficients δ_i diminishes as γ increases. According to Liu and Gertler (1987), there is a limiting value $\delta_x > 0$ so that the closed-loop is guaranteed to be stable for $|\delta_i| < \delta_x$. Therefore, a limiting value γ_m exists such that eqn. 2.94 is stable.

□□□

Theoretically, the minimum variance controller in eqn. 2.73 attempts to rapidly set the prediction to the setpoint in one step after a delay of k-steps. In practice this controller is not commonly employed in process control applications because it is extremely sensitive to variations in model parameters and non-minimum phase process.

The resulting performance in many cases is oscillatory responses and unacceptably vigorous control action, which requires further detuning strategies. On the other hand, a mean-level controller is known to be robust for open-loop stable processes, but slow in responses. Both controllers require only a unity control horizon which is sufficient for most applications. Therefore, γ becomes a weighting term to select either mean level control by allowing γ to increase or minimum variance by setting $\gamma = 0$. A nominal value gives the flexibility of trading off the rapid minimum variance with robust mean level control (*i.e.* trading off performance with robustness). The guidelines for selection of γ are described in the next chapter.

2.5 Conclusions

A long-range predictive control law with a terminal matching condition has been developed by minimizing the prediction errors over a finite horizon and at steady-state. In contrast to ordinary LRPC algorithms which require output predictions over a large prediction horizon, this modified control law only looks at the future outputs over a short prediction horizon as well as the infinite-time or steady-state prediction of the output. It implies that less computation and knowledge is required without losing the long range property. Both transfer function and convolution modeling approaches can benefit from this approach of using steady-state error weighting. The convolution modeling approach is also shown to be able to approximate a long-range predictive controller by knowing only the first few step response coefficients and the process gain instead of a set of up to 50 step response coefficients. The role of this weighting has been investigated in closed-loop control. It is shown to be more advantageous over ordinary control weighting because it not only causes zero offset at steady-state even when a model gain is in error, but also eliminates the additional closed-loop pole induced by the integrator from the disturbance modeling term. It is similar to the output prediction horizon, providing stabilizing effect for open-loop stable systems in the presence of modeling error and non-minimum zeros.

Bibliography

Auslander, D.M., Y. Takahashi, and M. Tomizuka. Direct digital process control: Practice and algorithms for microprocessor application. *Proc. IEEE*, 66(2):199–208, February 1978.

Bamberger, W. and R. Isermann. Adaptive on-line steady-state optimization of slow dynamic processes. *Automatica*, 14:223–230, 1978.

Bhattacharya, A. and B. Joseph. On-line optimization of chemical process. In *Proc. American Control Conference*, pages 334–337, Paper MP5, 1982.

Chesna, S.A. and B.E. Ydstie. Self tuning and adaptive control of chemical processes. In K.S. Narendra, editor, *Adaptive and learning systems: Theory and applications*, pages 119–130. Plenum Press, New York, 1986.

Clarke, D.W. and P.J. Gawthrop. Self-tuning controller. *IEE Proc.-D*, 122(9):929–934, 1975.

Clarke, D.W. and C. Mohtadi. Properties of Generalized Predictive Control. *Automatica*, 25(6):859–875, 1989.

Cutler, C.R. and B.L. Ramaker. Dynamic matrix control — a computer control algorithm. In *Proc. JACC*, San Francisco, 1980.

Garcia, C.E. and M. Morari. Optimal operation of integrated processing systems: Part I Open-loop on-line optimizing control. *A.I.Ch.E. Journal*, 27(6):960–968, 1981.

Garcia, C.E. and M. Morari. Internal model control. 1. A unifying review and some new results. *Ind. Eng. Chem. Process Des. Dev.*, 21:308–323, 1982.

Garcia, C.E. and M. Morari. Optimal operation of integrated processing systems: Part II Closed-loop on-line optimizing control. *A.I.Ch.E. Journal*, 30(2):226–234, 1984.

Garcia, C.E., D.M. Prett, and M. Morari. Model predictive control: Theory and practice - A survey. *Automatica*, 25:335–348, 1989.

Åström , K.J. and B. Wittenmark. On self tuning regulators. *Automatica*, 9:185–199, 1973.

Lambert, E.P. *Process Control Applications of Long Range Prediction*. D.Phil. thesis, University of Oxford, 1987.

Li, S., K.Y. Lim, and D.G. Fisher. A state space formulation for model predictive control. *A.I.Ch.E. Journal*, 35(2):241–249, 1989.

Lim, K.Y. Multivariable optimal constrained control algorithm (MOCCA). M.Sc. thesis, University of Alberta, 1988.

Liu, K. and J. Gertler. On-line stabilization of adaptive controllers by detuning in a supervisory framework. In *Proc. American Control Conference*, pages 194–200, Minneapolis, MN, 1987.

McFarlane, R.C. and D.W. Bacon. Adaptive optimizing control of multivariable constrained chemical processes. 1. Theoretical development. *Ind. Eng. Chem. Res.*, 28:1828–1834, 1989.

McIntosh, A.R., S.L. Shah, and D.G. Fisher. Analysis and tuning of adaptive generalized predictive control. *Canadian Journal of Chemical Engineering*, 69:97–110, February 1991.

Papadoulis, A.V., C.A. Tsiligiannis, and S.A. Svoronos. A cautious self-tuning controller for chemical processes. *A.I.Ch.E. Journal*, 33(3):401–409, 1987.

Richalet, J., A. Rault, J.L. Testud, and J. Papon. Model predictive heuristic control: applications to industrial processes. *Automatica*, 14:413–428, 1978.

Shook, D.S., C. Mohtadi, and S.L. Shah. A control-relevant identification strategy for GPC. *IEEE Trans. Auto. Cont.*, 1991. In press.

Sripada, N.R. and D.G. Fisher. Multivariable optimal constrained control algorithm (MOCCA): Part 1. Formulation and Application. In *Proc. International Conference on Industrial Process Modeling and Control*, volume 1, Hangzhou, China, June 1985.

Takahashi, Y., M. Tomizuka, and D.M. Auslander. Simple discrete control of industrial processes (Finite time setting control algorithm for single-loop digital controller). *Trans. ASME*, pages 354–361, December 1975.

Chapter 3

Evaluation of LRPC with a Terminal Matching Condition

3.1 Introduction

Multistep predictive controllers have been shown to be successful in many industrial applications. Early control strategies based on long-range multistep minimization methods usually required a known step or impulse response model for making future predictions (Richalet *et al.*, 1978; Cutler and Ramaker, 1980). These model-based predictive control algorithms such as MAC and DMC can be analyzed and investigated by considering the unifying theoretical framework of the IMC techniques (Garcia and Morari, 1982) based on the utilization of a "generic" internal model to predict the effect of manipulated variables on the output. Since on-line identification of large number of dynamics response coefficients is not efficient, some long-range predictive controllers (Peterka, 1984; Mosca *et al.*, 1984; de Keyser and van Cauwenberghe, 1985) make use of a low-order parametric models from which impulse or step response models can be derived. Generalized Predictive Control (GPC) (Clarke *et al.*, 1987) is another member of the class of long-range predictive controllers and a natural extension of the Generalized Minimum Variance controller (Clarke and Gawthrop, 1979) from single-step to multistep minimization. Its predictive structure is based on auto-regressive integrated moving-average model with an exogenous term.

All of these long-range control schemes basically minimize the squares of the errors between predicted process outputs and setpoints over a future prediction horizon. The extension of this approach to include a terminal matching condition has been developed in the preceding chapter. It is achieved by combining a long-range predictive control objective with a steady-state error weighting (γ-weighting) term so that the process response both within the prediction horizon and at steady-state are considered. As the controller becomes more sophisticated, deciding the best values for the tuning parameters for a system with little *a priori* process information may be difficult especially during the commissioning stage. The present work provides two tuning strategies for implementation of such techniques for commissioning chemical processes and evaluates the performance of γ-weighting term as a tuning parameter. Although various valuable guidelines have been suggested (Garcia and Morari, 1982; Marchetti *et al.*, 1983; Mohtadi and Clarke, 1986; Maurath *et al.*, 1988; Xi, 1989; Scattolini and Bittanti, 1990), this work follows the tuning philosophy of McIntosh *et al.* (1991) by giving specific recommendations for selecting tuning parameters. Evaluation of γ is performed by making comparisons between γ-weighting and output prediction horizon as well as ordinary control weighting. Experimental results obtained from controlling the temperature of a continuously stirred heating system by this new approach are also included in this study.

3.2 Control Algorithm Review

Steady-state error weighting (γ-weighting) stems from the idea that a long-range predictive controller can be realized by knowing the initial and terminal values of the process response. Section 2.2 describes different modeling approaches. Among them is a "limited" convolution model which includes n initial impulse response coefficients, h_i, and a steady-state gain, g_s. It can be transformed into a transfer function model by assuming a S-shaped low-frequency dynamics, *i.e.*

$$\begin{aligned} G_p &= h_1 q^{-1} + h_2 q^{-2} + \cdots + h_{n-1} q^{-n+1} + \frac{h_n q^{-n}}{1 - p q^{-1}} && (3.1) \\ &= \frac{b'_1 q^{-1} + b'_2 q^{-2} + \cdots + b'_n q^{-n}}{1 - p q^{-1}} && (3.2) \end{aligned}$$

where

$$\begin{aligned} b'_1 &= h_1 \\ b'_i &= h_i - p h_{i-1} \quad \text{for} \quad i = 1 \;\; to \;\; n \\ p &= 1 - \frac{h_n}{g_s - \sum_{i=1}^{n-1} h_i} \end{aligned}$$

Then predictor based on an ARIMAX model,

$$\hat{A} y(t) = q^{-k} \hat{B} u(t) + T \frac{\xi(t)}{\Delta} \tag{3.3}$$

is expanded to make future predictions for a short prediction horizon,

$$\hat{y}(t + j \mid t) = G_j \Delta u(t + j - k) + H_j \frac{\Delta u(t - k)}{T} + F_j \frac{y(t)}{T} \tag{3.4}$$

and at steady state,

$$\hat{y}(s|t) = g_s \sum_{j=1}^{nu} \Delta u(t + j - k) + H_s \frac{\Delta u(t - k)}{T} + F_s \frac{y(t)}{T} \tag{3.5}$$

where G, F and H are derived from the following diophantine identities:

$$\begin{aligned} T &= E_i \hat{A} \Delta + q^{-i} F_i && (3.6) \\ E_i \hat{B} &= G_i T + H_j q^{-i} && (3.7) \end{aligned}$$

and "s" in eqn. 3.5 denotes a value at steady state.

The performance index is an extension of GPC to include a steady-state error weighting:

$$\begin{aligned} J &= \sum_{j=n1}^{n2} \left[\hat{y}(t+j) - w(t+j)\right]^2 + \sum_{j=1}^{nu} \lambda(j) \left[\Delta u(t+j-1)\right]^2 \\ &\quad + \sum_{j=1}^{nu} \gamma(j) \left[\hat{y}(s|t+j-1) - w(s)\right]^2 \end{aligned} \tag{3.8}$$

The control law minimizing J is found to be

$$\mathbf{U} = \left[\mathbf{G}^T\mathbf{G} + \Lambda + \mathbf{G_s}^T\Gamma\mathbf{G_s}\right]^{-1} \left[\mathbf{G}^T(\mathbf{W} - \mathbf{f}) + \mathbf{G_s}^T\Gamma(\mathbf{W_s} - \mathbf{f_s})\right] \tag{3.9}$$

where

$$\begin{aligned}
\mathbf{U} &= \left[\ \Delta u(t+n1-1) \quad \Delta u(t+n1) \quad \cdots \quad \Delta u(t+n2-1)\ \right]^T \\
\mathbf{G} &= \begin{bmatrix} g_{n1-1} & \cdots & g_0 & 0 & 0 & \cdots & 0 \\ g_{n1} & \cdots & g_1 & g_0 & 0 & \cdots & 0 \\ \vdots & \ddots & & & \ddots & & \vdots \\ \vdots & & \ddots & & & & g_0 \\ \vdots & \cdots & & & & & \vdots \\ g_{n2-1} & g_{n2-2} & \cdots & & & & g_{n2-nu} \end{bmatrix} \\
\mathbf{f} &= \left[\ f(t+n1) \quad f(t+n1+1) \quad \cdots \quad f(t+n2)\ \right]^T \\
f(t+j) &= H_j \,\frac{\Delta u(t-1)}{T} + F_j \,\frac{y(t)}{T} \\
\mathbf{G_s}\ (nu \text{ x } nu) &= \begin{bmatrix} g_s & 0 & \cdots & 0 \\ g_s & g_s & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ g_s & \cdots & \cdots & g_s \end{bmatrix} \\
\mathbf{f_s}\ (nu \text{ x } 1) &= \left[\ 1 \quad 1 \quad \cdots \quad 1\ \right]^T \cdot \left(H_s \,\frac{\Delta u(t-k)}{T} + F_s \,\frac{y(t)}{T}\right) \\
\Lambda &= \operatorname{diag}\left[\lambda(1) \quad \lambda(2) \quad \cdots \quad \lambda(nu)\right] \\
\Gamma &= \operatorname{diag}\left[\gamma(1) \quad \gamma(2) \quad \cdots \quad \gamma(nu)\right] \\
\mathbf{W} &= \left[w(t+n1) \quad w(t+n1+1) \quad \cdots \quad w(t+n2)\right]^T
\end{aligned}$$

$$\mathbf{W}_s\ (nu \times 1) \;=\; [1\ \ 1\ \ \cdots\ \ 1]^T \cdot w(s)$$

For adaptive control with a receding horizon policy (Ydstie, 1984), only the first element of **U** is implemented. A new **U** is calculated at the next sampling period.

3.3 Performance Tuning of γ-Weighting

The fundamental tuning parameters for GPC with steady-state error weighting are listed below:

- minimum output horizon $n1$
- maximum output horizon $n2$
- control horizon nu
- steady-state error weighting γ

Since GPC without γ-weighting can be reduced to many of the well-known control methods by different selections of the first three tuning parameters, two new tuning strategies are recommended by incorporating γ as an "active" parameter into two common control methods: deadbeat and mean-level control. While keeping the other tuning parameters constant, γ is adjusted to obtain the desired response. Depending on the performance required, the first of the following approaches can be used for fast response whereas the second one for robust operation.

3.3.1 Strategy #1: Deadbeat Approach

Clarke and Mohtadi have shown that the equivalent of a stable deadbeat controller for an observable and controllable linear system is obtained by the following selections:

$$nu = N, \quad n1 = N, \quad n2 \geq 2N - 1, \lambda = 0 \tag{3.10}$$

where N is the order of a state space model (Mohtadi, 1987). When an input/output model is considered, the configuration of a deadbeat controller requires the following parameter settings:

$$nu = na + 1, \quad n1 = nb + 1, \quad n2 = nu + n1 - 1 \tag{3.11}$$

The settings specified above are, therefore, fixed at the commissioning stage. Then γ is adjusted to tailor the speed of response.

In principle, the value of γ can vary from zero to infinity. A very large value of γ only makes the controller behave more like a mean-level controller. If the magnitude of the matrix inverse is approximated by the trace of $\mathbf{G}^T\mathbf{G}$ as in McIntosh's tuning strategy (1991), γ can be selected according to the following "formula".

$$\gamma = \frac{2 \cdot m \cdot \mathrm{tr}[\mathbf{G}^T\mathbf{G}]}{(nu + 1) \cdot nu \cdot g_s^2} \tag{3.12}$$

where m is the degree of detuning from deadbeat to mean-level. The control increments are roughly the sum of deadbeat output response weighted by $\frac{1}{m+1}$ plus the mean-level response weighted by $\frac{m}{m+1}$. The initial value of γ can be chosen by setting $m = 1$ so that the control performance is about the average of the two responses for commissioning.

Because the formulation of γ-weighting already has g_s^2 as a factor in the inverse portion of eqn. 3.9, the closed-loop performance is independent of any gain changes. This is an advantage of γ over the use of λ-weighting which often has to be adjusted relative to g_s^2 by scaling. Another advantage is that the slowest control response resulting from a large value of γ is limited to that of mean-level control whereas increasing λ can indefinitely decreases the response.

3.3.2 Strategy #2: Mean-Level Approach

For typical industrial plant models, a large value of $n2$ and a value of nu of 1 are usually sufficient to produce satisfactory control performance for open-loop stable plants. The control law without any weighting on control movements can be written simply as:

$$\Delta u = \frac{\sum_{i=n1}^{n2} g_{i-1}\,(w - f(t+i))}{\sum_{i=n1}^{n2} g_{i-1}^2} \tag{3.13}$$

It is obvious from the equation that inverse model effect should be avoided by choosing a large $n2$ such that $\sum_{i=n1}^{n2} g_{i-1}$ is strictly positive. Although a default output horizon window of $n2 - n1 + 1 = 10$ has been found to produce robust response, a larger $n2$ corresponding to a 90 % rise-time of the process is more desirable especially when commissioning a new controller. Depending on the sampling time, the size of $n2$ capturing the rise-time can vary from 20 to 50, adding a lot of computational load to the controller.

When γ-weighting is used in conjunction with $n2$, it becomes the choice for performance tuning. Then the value of $n2$ can be reduced considerably to capture only the high frequency dynamics which usually appear during the first "time constant" period. This is similar to the guideline suggested by Maurath *et al.*(1985) recommending $n2$ equivalent to the number of sampling periods required for the process open-loop step response to reach 50 % of its final value. Therefore, the configuration for commissioning a process is proposed to be

$$n1 = k, \quad n2 = \frac{\tau}{\tau_s} + k, \quad nu = 1, \quad \sum_{i=n1}^{ns} g_{i-1} + \gamma > 0 \tag{3.14}$$

where

- τ is the dominant or overall time constant of the process

- τ_s is the sampling time

For example, if the sampling period is chosen by taking one-fifth of a time constant, $n2$ will be $5 + k$ where k is the units of time delay. Starting with a large value of γ, the process response can be adjusted by gradually decreasing γ until satisfactory results are observed. The control law for this approach is given below:

$$\Delta u = \frac{\displaystyle\sum_{i=n1}^{ns} g_{i-1}\left(w(t+i) - f(t+i)\right) + g_s\gamma\left(w(s) - f(s)\right)}{\displaystyle\sum_{i=n1}^{ns} g_{i-1}^2 + g_s^2\gamma} \tag{3.15}$$

where ns denotes the reduced output prediction horizon in order to distinguish it from $n2$ in eqn. 3.13.

One can obtain a feel for the range of admissible γ values by equating eqns. 3.13 and 3.15 and finding the values of γ corresponding to different $n2$'s. Without any loss of generality, the set point in both equations are assumed zero. After some re-arrangement, the resulting relationship between γ and $n2$ is given as:

$$\gamma = \frac{\displaystyle\sum_{i=n1}^{ns} g_{i-1}f(t+i) \sum_{i=ns+1}^{n2} g_{i-1}^2 - \sum_{i=ns+1}^{n2} g_{i-1}f(t+i) \sum_{i=n1}^{ns} g_{i-1}^2}{\displaystyle k^2 \sum_{i=n1}^{n2} g_{i-1}f(t+i) - kf_s \sum_{i=n1}^{n2} g_{i-1}^2} \tag{3.16}$$

It should be noted that the value of γ in eqn. 3.16 is dependent neither on the process gain nor the process time constant (τ), but only on the choice of $n1$, $n2$, ns, and the ratio $\frac{\tau}{\tau_s}$.

In order to show that γ is independent of the process gain, the gain term is factored out from all step response coefficients, *i.e.*

$$g_j = g_s g_j'$$

where the prime denotes a quantity corresponding to the case where the gain is unity. With this substitution, all gain terms in eqn. 3.16 will be canceled out. What remains is a relation between γ and a combination of the model response coefficients with a standardized unity gain.

$$\gamma' = \frac{\sum_{i=n1}^{ns} g'_{i-1} f'(t+i) \sum_{i=ns+1}^{n2} (g'_{i-1})^2 - \sum_{i=ns+1}^{n2} g'_{i-1} f'(t+i) \sum_{i=n1}^{ns} (g'_{i-1})^2}{\sum_{i=n1}^{n2} g'_{i-1} f'(t+i) - f'_s \sum_{i=n1}^{n2} (g'_{i-1})^2} \tag{3.17}$$

It should be noted again that the choice of γ is dependent on the discretization factor, $\frac{\tau}{\tau_s}$, but not on τ. The reason is that when two equal-gain and equal-order processes are discretized by the same factor of the individual time constants, *i.e.* $\frac{\tau 1}{\tau 1_s} = \frac{\tau 2}{\tau 2_s}$, the resulting step response models are exactly equal. Therefore, an approximate value of γ and ns which will provide an effect similar to a larger $n2$ value (with $\gamma = 0$) can be computed from eqn. 3.17 when $n1$ and $\frac{\tau}{\tau_s}$ are given. Fig. 3.1 shows the trajectories of γ corresponding to different $n2$ and ns. For this particular plot, $n1$ is one and the sampling time is taken as one-tenth of the time constant which is considered to be an appropriate discretization interval for most industrial processes.

3.4 Demonstration of γ-Weighting

The performance using γ-weighting is compared with λ and $n2$ in this section. Table 3.1 shows the details of three open-loop stable transfer functions with their poles and zeros. Process A is a second-order underdamped system with a damping factor of 0.5. Process B which has been used by McIntosh (1988) is a non-minimum phase system. Simulations for model-plant-mismatch is achieved by approximating the third-order Rohrs' model with a reduced-order model. Experimental results are obtained by controlling the effluent temperature of a continuous-stirred tank heater (CSTH).

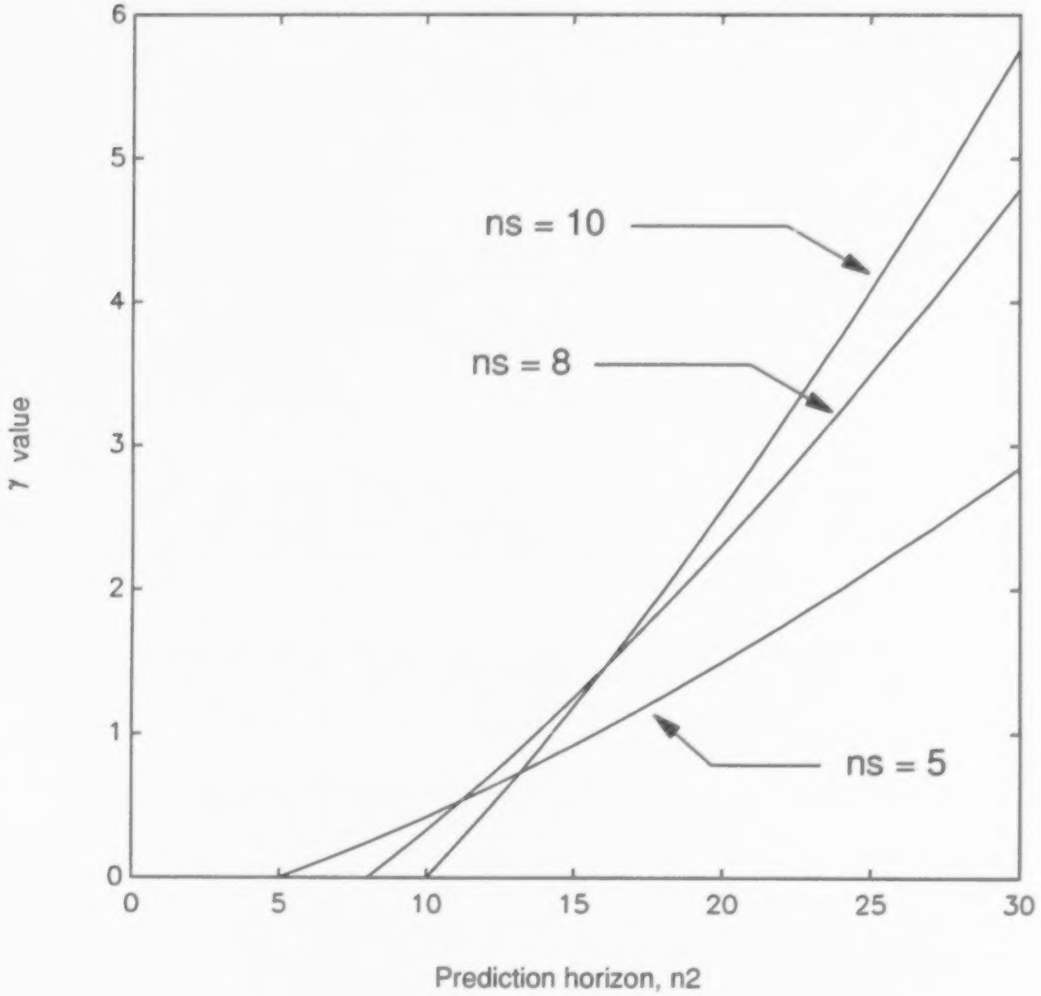

Figure 3.1: Trajectories of γ for different values of $n2$ for the special case of $nu = 1$

Process	Continuous Time	Discrete Time	Poles	Zeros
A	$\frac{1}{s^2+.5s+1}$	$\frac{.113z^{-1}+.109z^{-2}}{1-1.562z^{-1}+.779z^{-2}}$ $\tau_s = .5$	$.7811\pm$ $.4107i$	$-.9194$
B	$\frac{-2s+1}{(6s+1)(2s+1)}$	$\frac{-.0864z^{-1}+.468z^{-2}}{1-1.453z^{-1}+.5134z^{-2}}$ $\tau_s = 1$	.6065 .8465	1.699
C	$\frac{2(229)}{(s+1)(s^2+30s+229)}$	$\frac{.037z^{-1}+.0717z^{-2}+.00785z^{-3}}{1-1.342z^{-1}+.446z^{-2}-.045z^{-3}}$ $\tau_s = .1$	.9048 $.2127\pm$ $.0443i$	$-.1164$ -1.822

Table 3.1: Process models for computer simulations

3.4.1 GPC with γ-Weighting

Process A

It has been shown in Section 2.4.2 that both γ-weighting and λ-weighting penalize the incremental control actions when an ARIMAX model is assumed. However, unlike λ which brings an integrator into the closed-loop, GPC with γ and without λ does not have the additional integrating pole in the closed-loop model. The characteristic polynomial with γ-weighting is similar to that when no integrator is present (*i.e.* ARMAX modeling). This property is illustrated by comparing the root loci of the following two closed-loop characteristic polynomials:

ARIMAX model with weighted one-step ahead control:

$$\mathcal{C}_1 = g_{k-1}B + \lambda \Delta A \tag{3.18}$$

ARIMAX model with one-step ahead control and steady-state error weighting

$$\mathcal{C}_2 = g_{k-1}B + \gamma g_s^2 A \tag{3.19}$$

where g_{k-1} is the first non-zero element in B.

Root loci of eqns. 3.18 and 3.19 for process model A are plotted in figs. 3.2 and 3.3 as a function of λ and γ respectively. Since one-step ahead control with

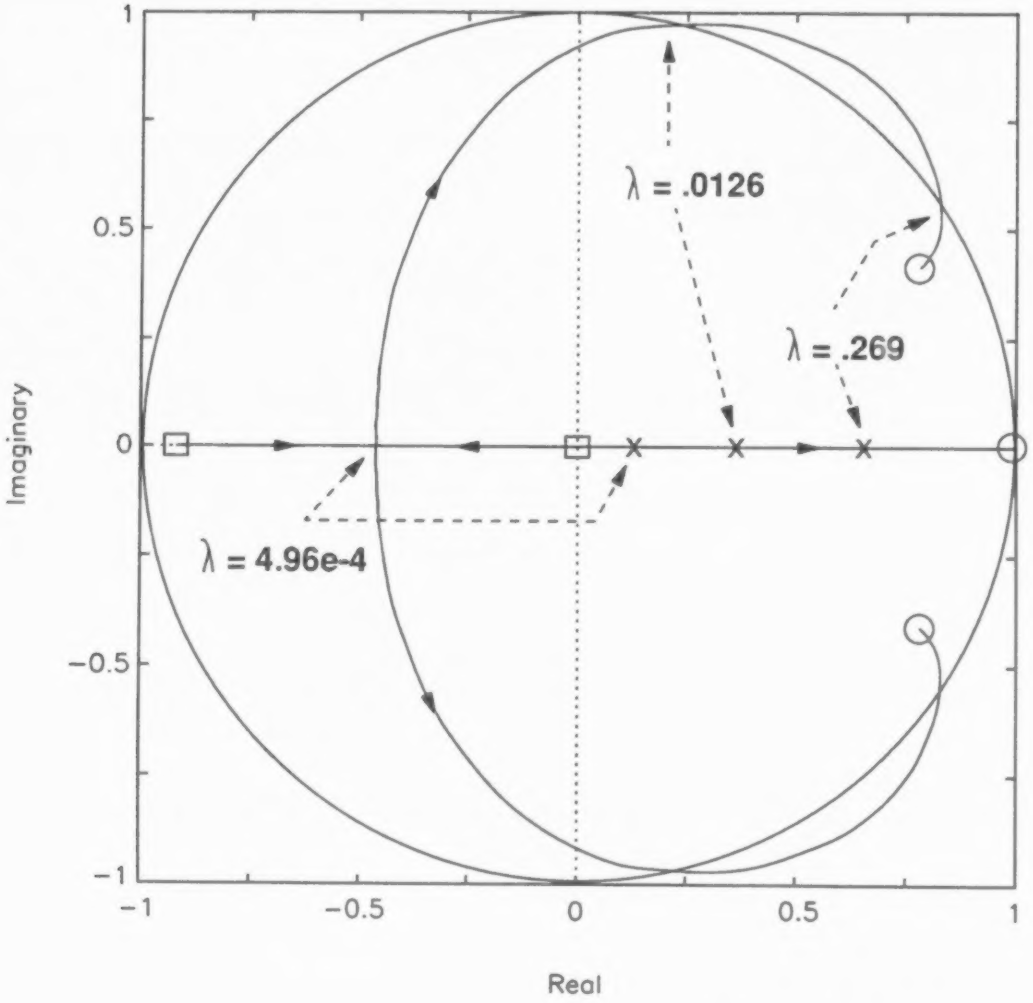

Figure 3.2: Closed-loop root locus of process A using minimum variance control with λ-weighting

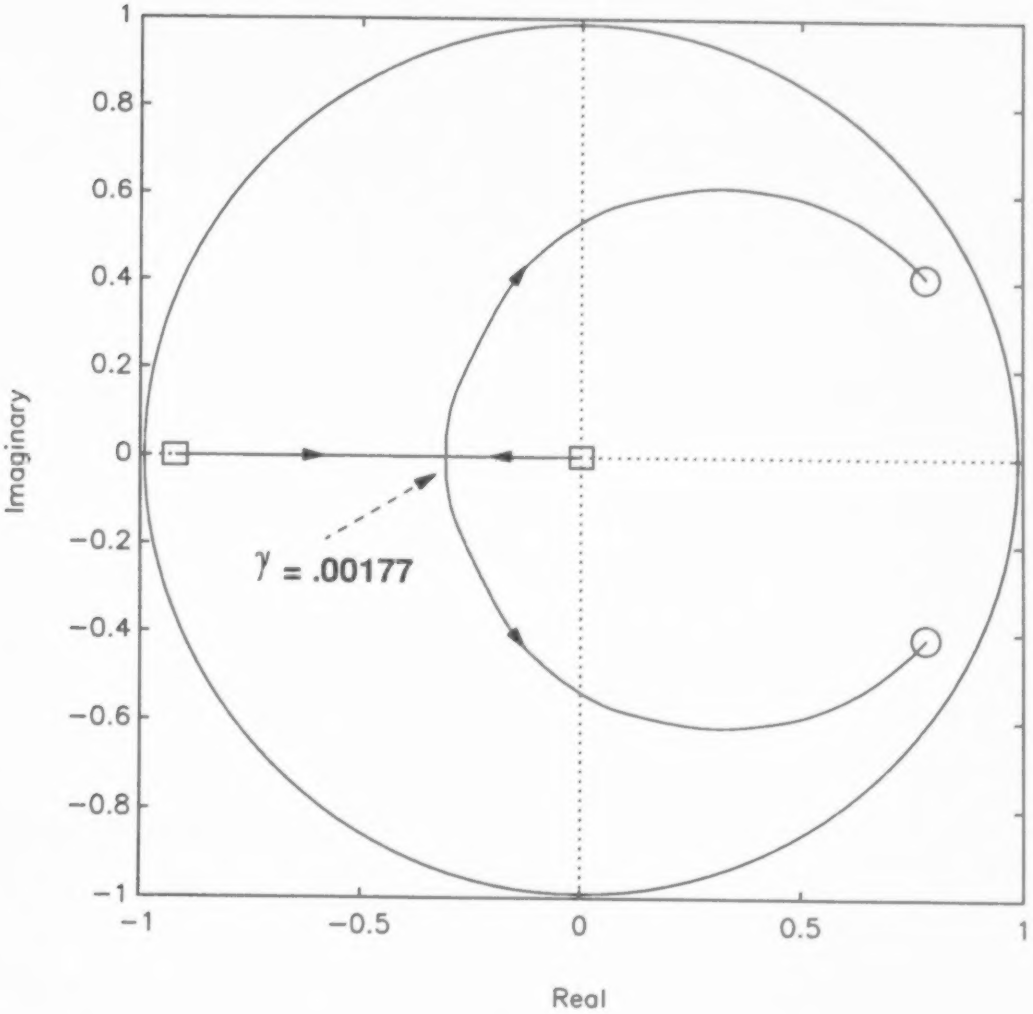

Figure 3.3: Closed-loop root locus of process A using minimum variance control with γ-weighting

$\lambda = 0$ is the same as having a minimum variance controller, the root locus starts at the open-loop zeros (represented by $\Box$ in the figure) and migrates to the open-loop poles (represented by $\circ$). In fig. 3.2, two roots move into the unstable region before converging to the open-loop poles while the additional root introduced by the integrator migrates outwards along the real axis to the unit circle as λ increases. The unstable model is due to the combination of detuned control and the inherent oscillatory nature of process A. Fig. 3.3 shows the effect of using γ-weighting. As γ increases, neither an unstable nor an additional order is observed. The advantage of no additional order when using γ-weighting is also true for any choice of $n1$ and $n2$ with $nu = 1$.

According to the tuning Strategy #1, process A is best controlled by a dead-beat controller while using γ as the "active" detuning parameter. Dead-beat control is achieved by setting $(n1, n2, nu) = (3, 5, 3)$.The closed-loop root loci at increasing γ is depicted in fig. 3.4. It is interesting to note that one pole always stays at the origin while the other two approach the open-loop poles. The overall closed-loop is indeed second order when $\gamma > 0$; therefore, $(n1, n2, nu) = (2, 4, 2)$ with γ-weighting is sufficient for controlling process A. The output performance is shown in fig. 3.5. The value of γ is calculated from the scaling formula 3.12 by setting $m = 0, .33, 1, 3$ for each upward set-point change. As m increases, the controller changes from dead-beat-dominant to mean-level-dominant. Therefore, the oscillatory response in fig. 3.5 (inherent in the underdamped open-loop model) becomes more apparent when $m = 3$.

Process B

A minimum of 5 output predictions from the non-minimum phase process B is required to bring all poles into the unit circle (see fig. 3.6) when $nu = 1$. The mean-level

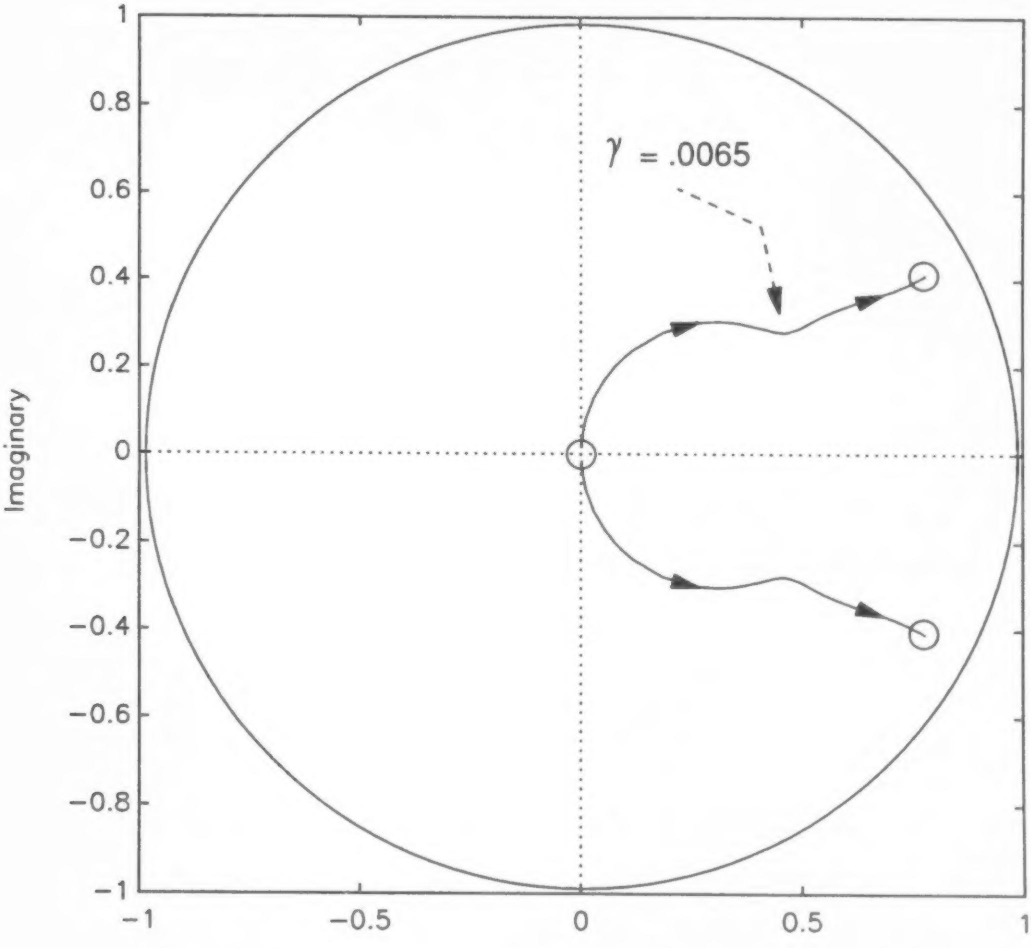

Figure 3.4: Closed-loop root locus of process A using dead-beat control with γ-weighting

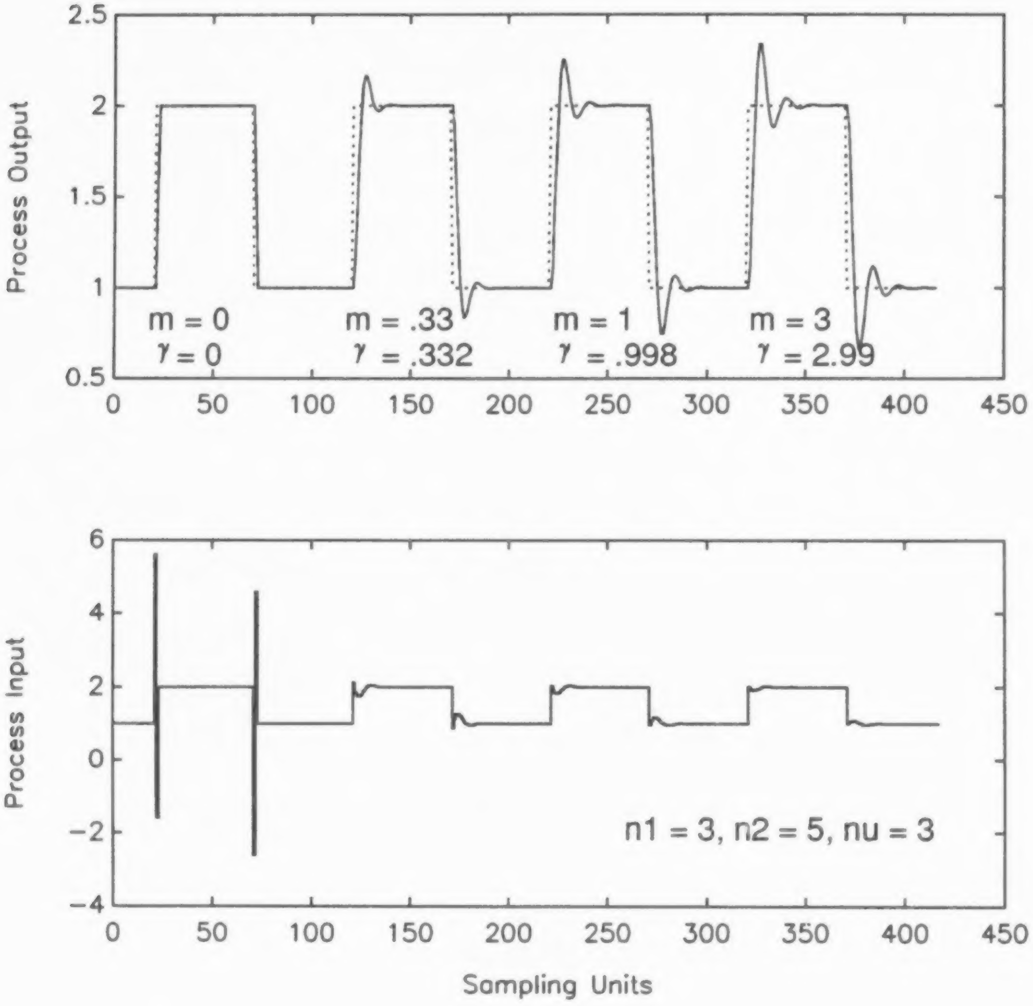

Figure 3.5: Closed-loop performance for process A using dead-beat control with γ-weighting

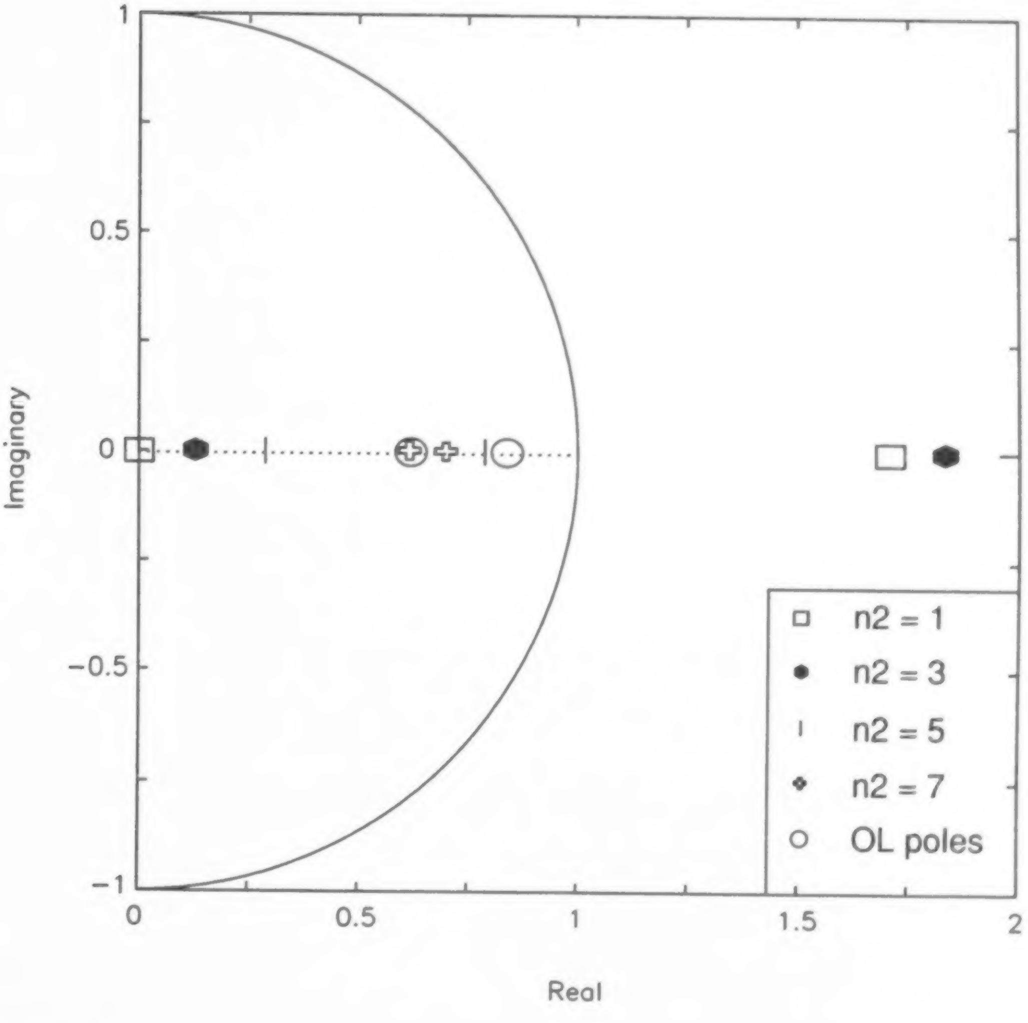

Figure 3.6: Closed-loop poles of process B using standard GPC

property of γ-weighting is compared to the infinite λ-weighting through their root loci shown in figs. 3.7 and 3.8. $(n1, n2, nu)$ is fixed at $(1, 5, 1)$. It is remarkable that by increasing γ the closed-loop poles are always real whereas complex roots are obtained by increasing λ. In fact, λ traverses three different regions listed in table 3.2.

Region	Detuning performance
$0 < \lambda < .00315$	ineffective
$.00315 \leq \lambda \leq 2.9$	slow output response, but oscillatory with overshoot
$\lambda > 2.9$	very slow response

Table 3.2: Effect of λ-weighting on Process B

The reason why γ behaves differently is due to its weighting on the steady-state prediction value such that the slowest output response is at most a mean-level result. Figs. 3.9 and 3.10 are used to illustrate the difference between the two weightings. In both figures, γ and λ are selected to be 0, 0.3, 1, and 3 covering all of the regions while $(n1, n2, nu) = (1, 5, 1)$. It is apparent that increasing γ results in mean-level control. However, first increasing λ to 1 causes small oscillation and overshoot; when $\lambda = 3$, the output is too slow to reach the setpoint within the time allowed.

It should be emphasized that process B can also be stabilized by increasing γ even if $n2$ is less than 5. However, a large value of λ will not necessarily stabilize a non-minimum phase system (Garcia and Morari, 1982).

Process C

Model-plant structural mismatch is introduced by using the "limited"-knowledge convolution model approach to describe process C. Recall from Section 3.2 that this approach requires only a small model horizon, n (*i.e.* the first n step response co-

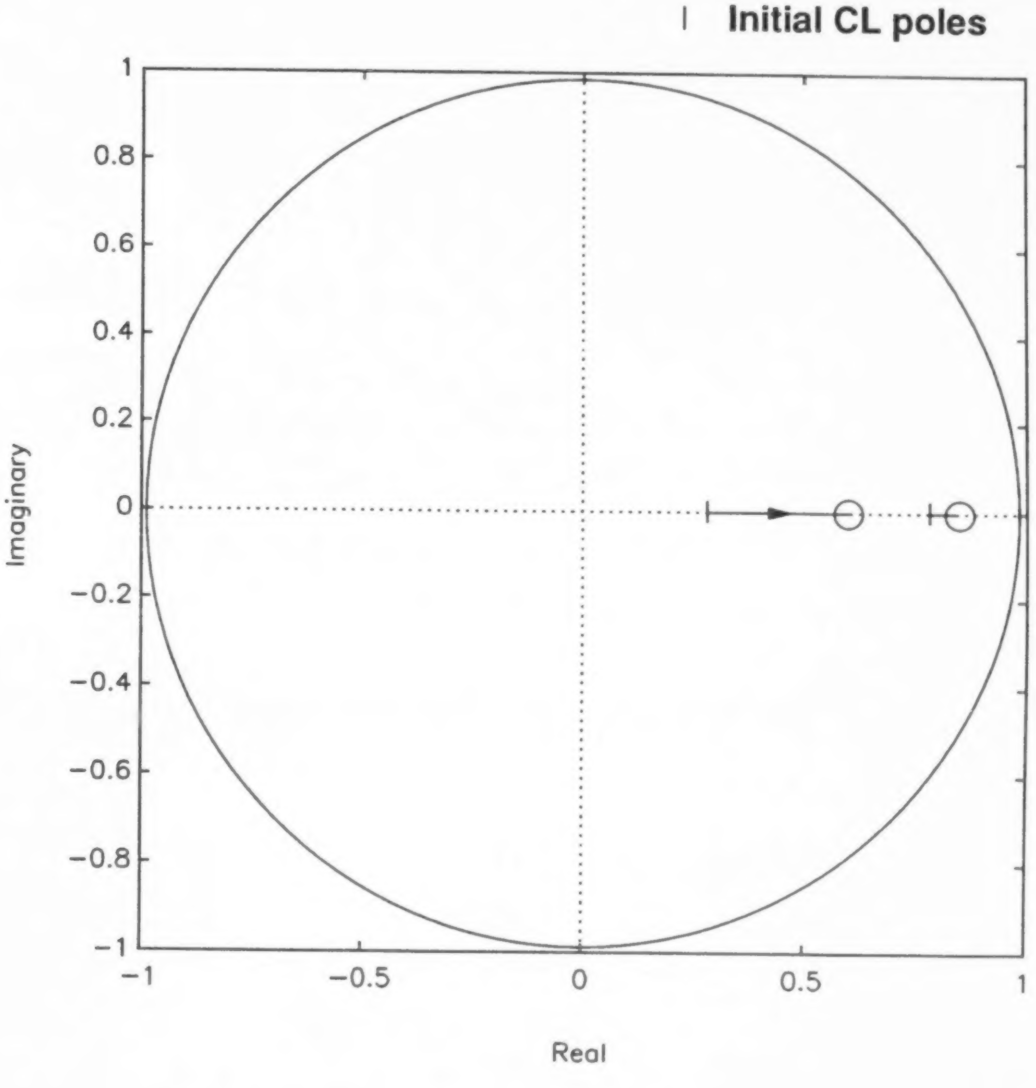

Figure 3.7: Closed-loop root locus of process B using GPC with γ-weighting

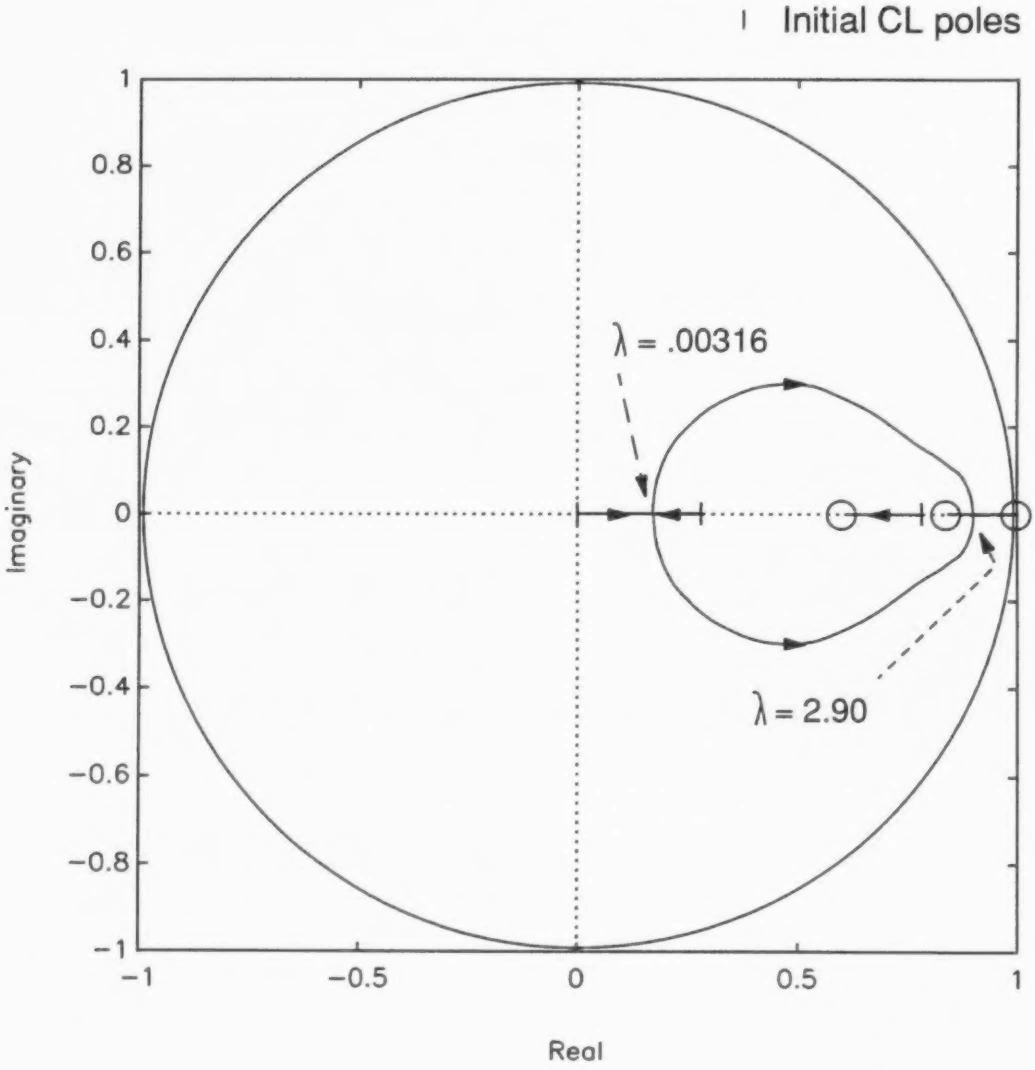

Figure 3.8: Closed-loop root locus of process B using GPC with λ-weighting

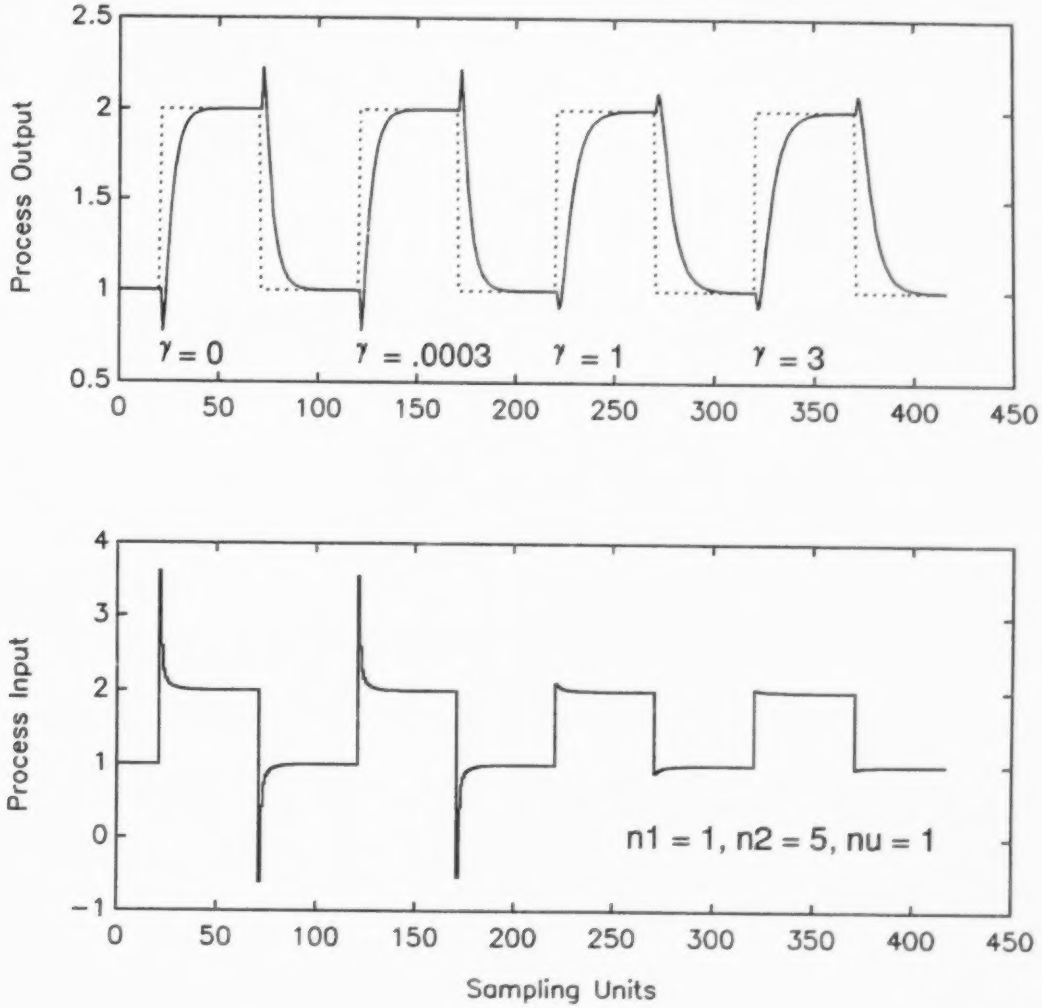

Figure 3.9: Closed-loop performance for process B using GPC with γ-weighting

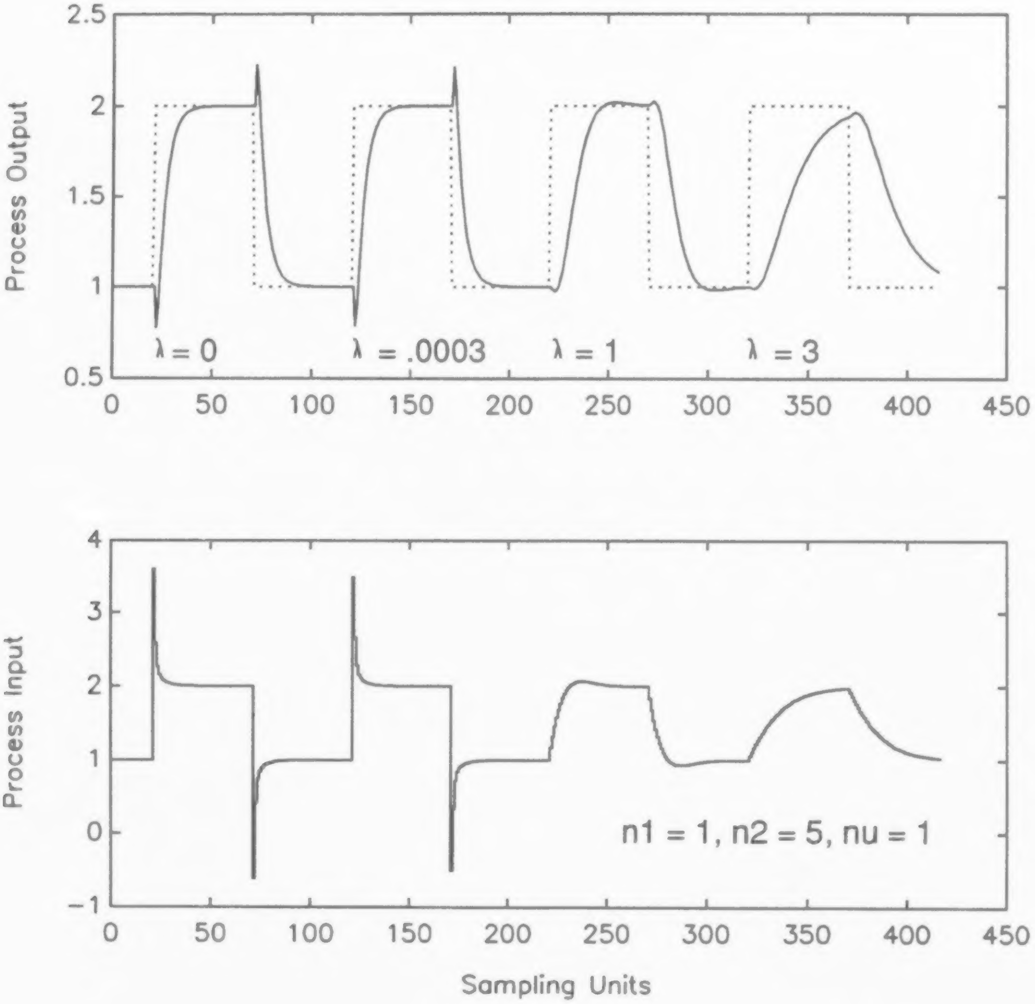

Figure 3.10: Closed-loop performance for process B using GPC with λ-weighting

efficients), and a known steady-state gain to approximate the process. Fig. 3.11 compares the step response trajectories from using different n with the true process step response model. When n is above 3, the trajectories almost coincide with the true one.

Fig. 3.12 shows the output performance of using $n = 3$ and gain of 2 (the true process gain) to approximate the true process. Because the initial high frequency dynamics is captured by the model, the controller performs favorably with γ-weighting even though $n2$ is only 5. The choices of γ shown in the figure are obtained from fig. 3.1, approximating a longer $n2$ of 10, 15, and 20, respectively.

3.4.2 Adaptive GPC with γ-Weighting

Process C

GPC with γ-weighting is made adaptive by estimating model parameters on-line by a recursive least squares algorithm (RLS) implemented with a variable forgetting factor. Process C is approximated by a first-order model which causes a considerable model-plant mismatch. Since Rohrs' model is predominantly first-order at low frequency with a high frequency second order dynamics, parameter estimation using ordinary RLS requires good filtering techniques (Shook *et al.*, 1991). Otherwise, performance would deteriorate or even become unstable due to inadequate modeling. McIntosh found that the parameters estimated by using a high pass filter $\frac{\Delta}{(1-0.9q^{-1})}$ are $\hat{a}_1 = -.94$ and $\hat{b}_0 = 0.033$ which gives a gain of 0.55 (McIntosh, 1988). The model gain obtained by RLS during dynamic simulation varied between 0.17 and 104. Because γ-weighting strongly depends on the model gain for steady-state predictions, a fixed model gain which assumes an error of 10 % below the true gain is provided for the controller. There are other methods to deal with the modeling problem; however, they are beyond

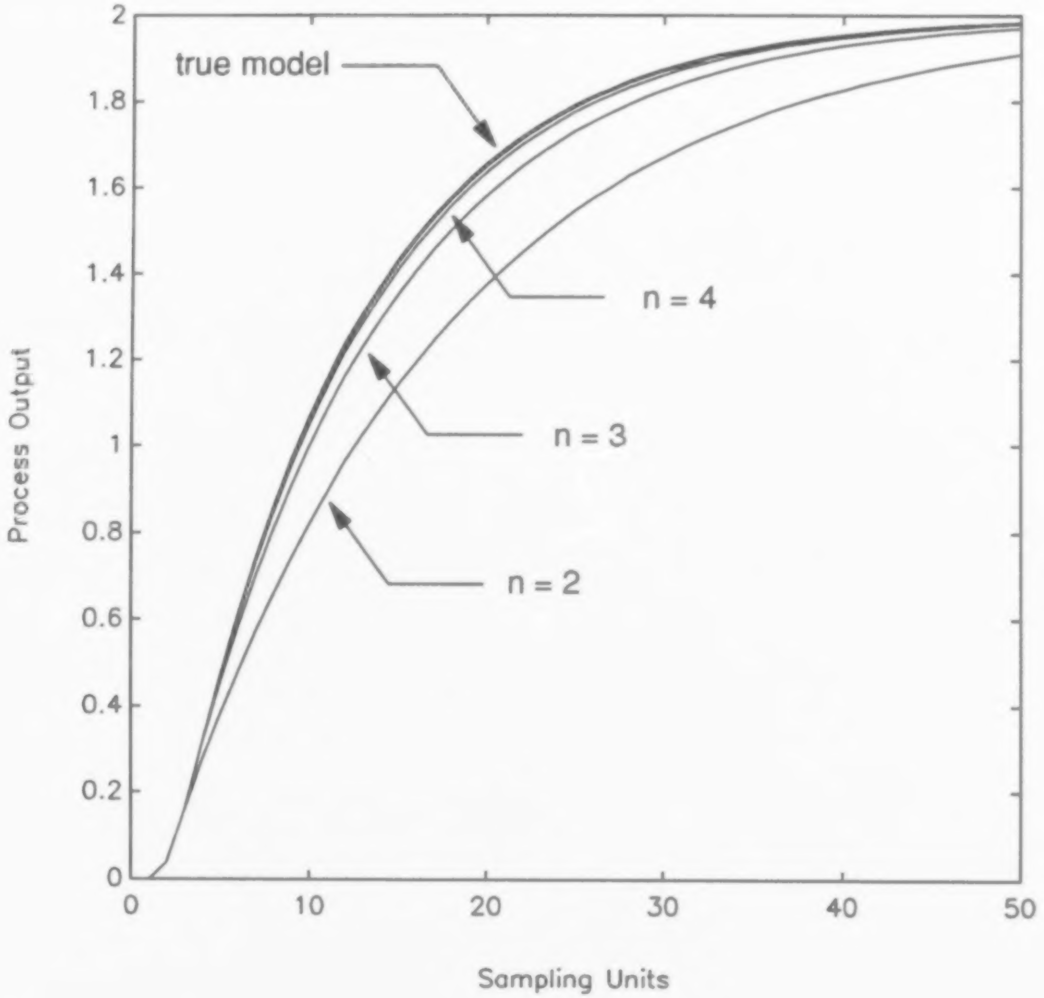

Figure 3.11: Step response trajectories of modeling process C using various modeling horizons

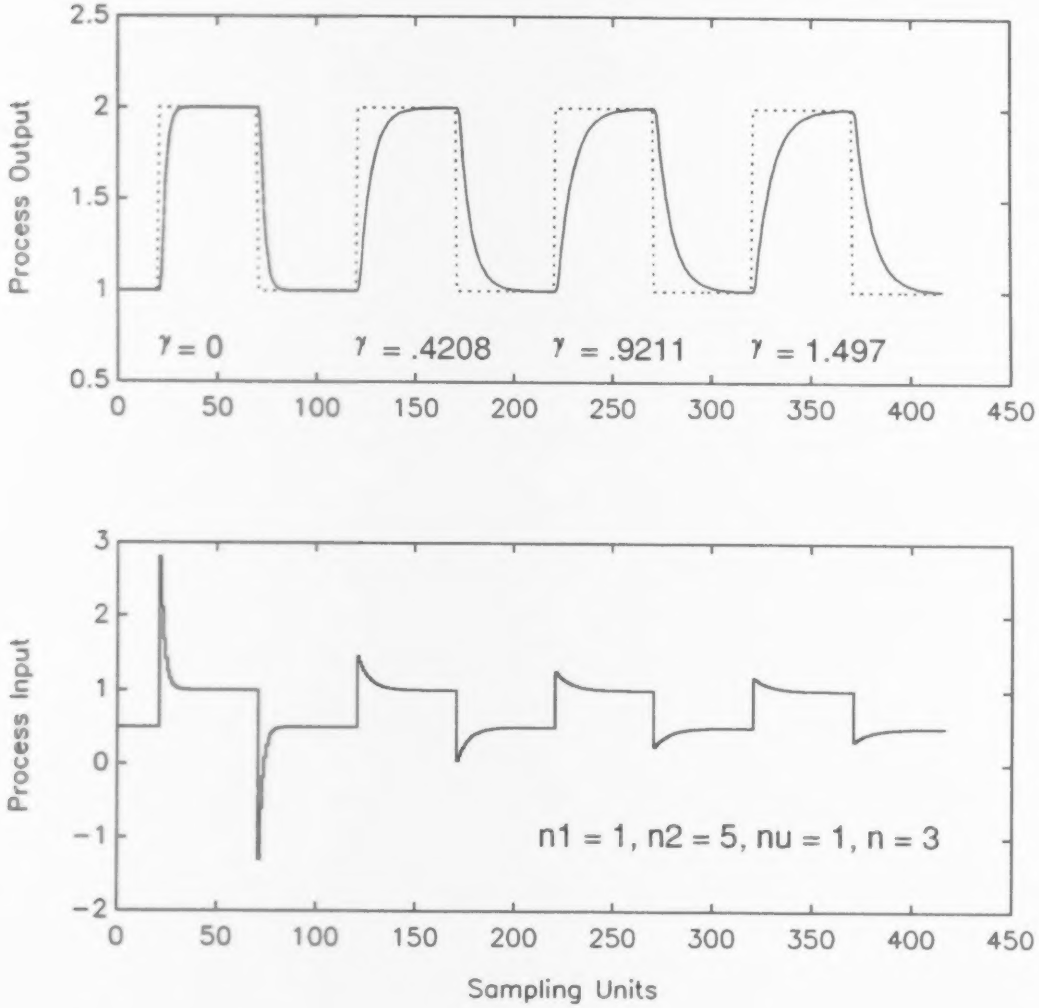

Figure 3.12: Closed-loop performance for process C using GPC with γ-weighting

the scope of this study and will not be discussed further.

Fig. 3.13 demonstrates the robustness of using γ-weighting. The simulation is initialized by a pseudo-random binary sequence (not shown in the graph) followed by four set-point changes. According to the tuning Strategy #2, γ is selected to be 2 to enhance robustness initially. Then it is decreased to tailor the process response. Although a 10 % offset between the model and process gain is present, the controller with γ-weighting and a short output prediction horizon ($n2 = 5$) manages to provide stable and satisfactory response.

Continuous Stirred-Tank Heater

Experimental demonstrations were carried out on a pilot-scale continuous-stirred tank in the Department of Chemical Engineering at the University of Alberta. Fig. 3.14 shows the schematic diagram of the equipment. It consists of a glass tank 50 cm high with an inside diameter of 14.5 cm. Cold water enters the tank from the top through a 1.5 cm-diameter pipe. After being heated by a steam coil, the process water leaves the tank and passes through a copper pipe which contains four thermocouples (located at different distances downstream of the tank to provide true delays) to measure the effluent temperature. For this demonstration, thermocouple #1 was chosen, introducing a delay of about one-fifth of a time constant. The water temperature was manipulated by a steam control valve and disturbed by increasing the amount of cold water entering the tank. A proportional DP-cell controller was used to maintain the water level constant. An opto-22 multiplexer provides A/D and D/A conversions, sending information to an IBM PS/2 Model 70 computer operating under the QNX operating system[1]. The computer provides an operator-program interface called Multicon (Qiu *et al.*, 1988) which performs real-time scheduling and

---

[1] QNX, Quantum Software Systems Ltd., 1988

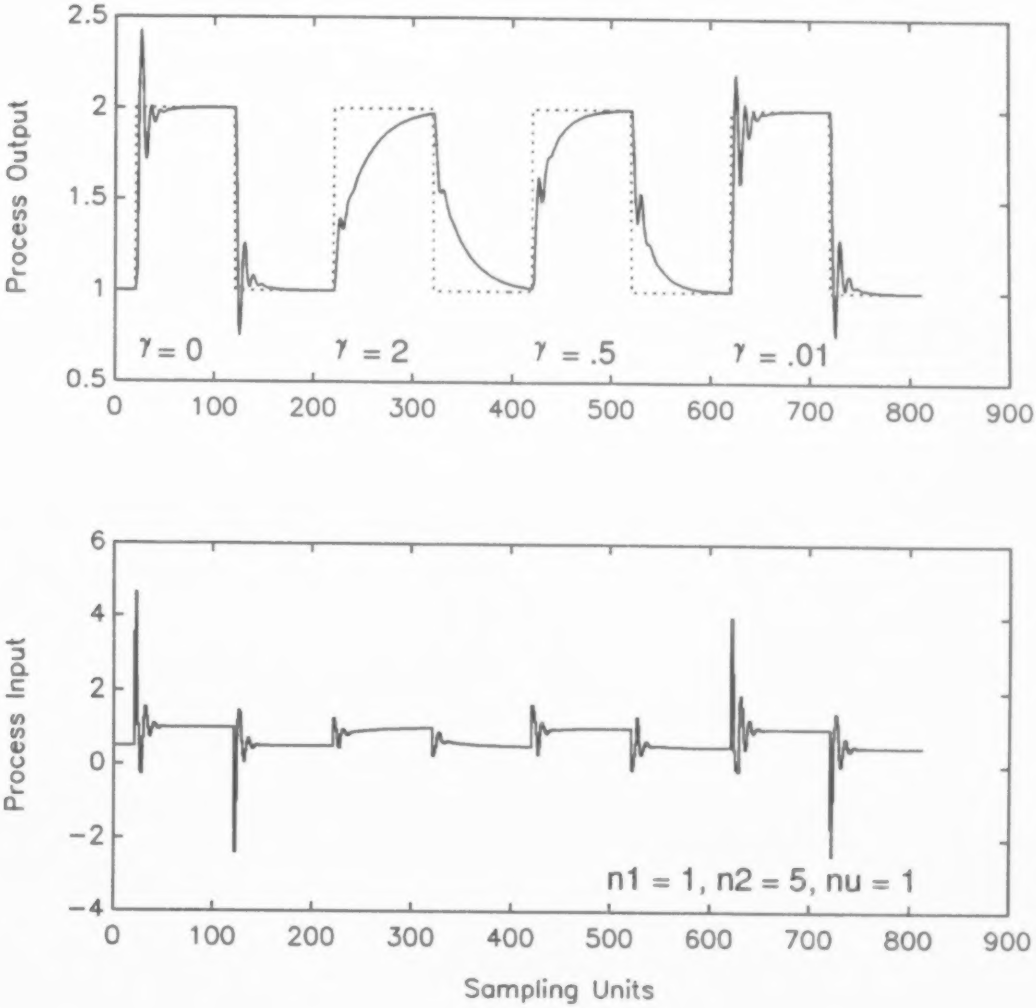

Figure 3.13: Closed-loop performance for process C using adaptive GPC with γ-weighting

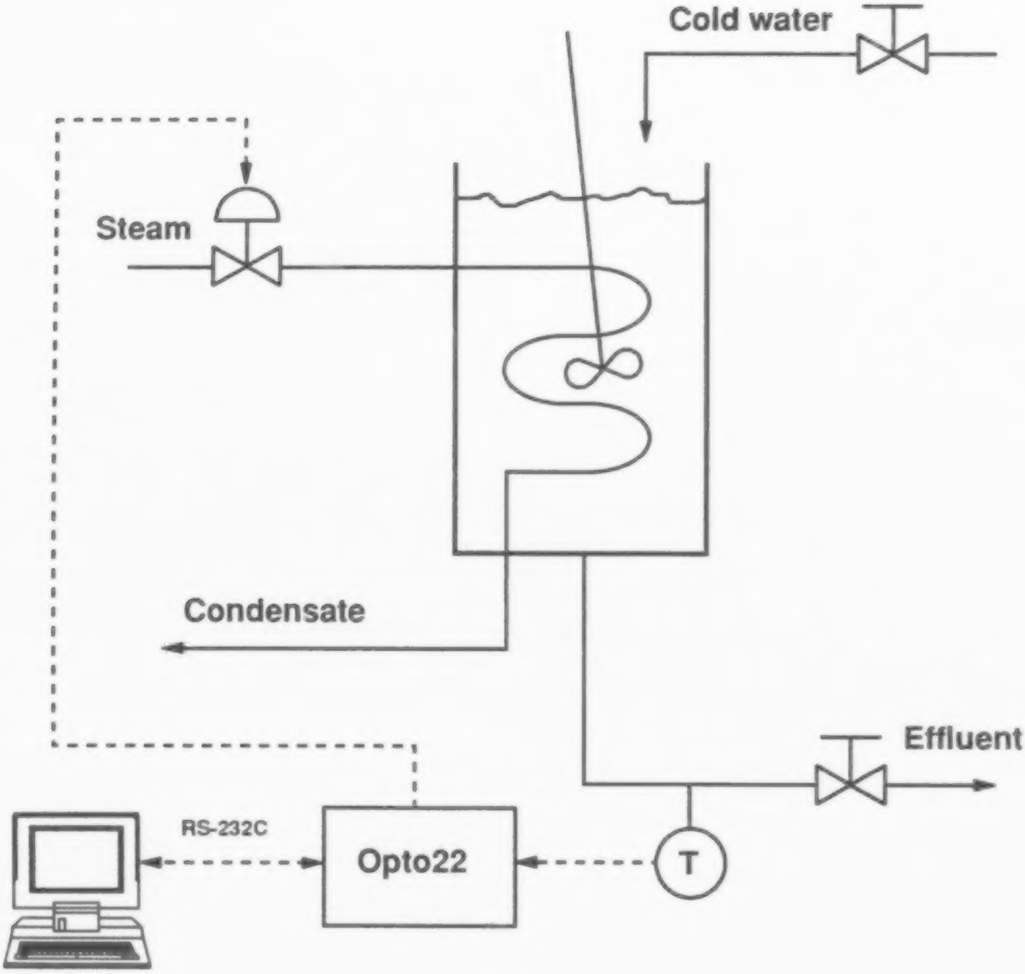
Cold water
Steam
Condensate
Effluent
RS-232C
Opto22
T

Figure 3.14: Schematic diagram of continuous stirred-tank heater

multi-tasking operations. All calculations and on-line changes of tuning parameters were done within Multicon.

Inlet water flow rate was set at .57 m/s. During the disturbance period, it was increased to .78 m/s. The sampling time was 5 seconds. The overall time constant is about one minute. Model parameters were estimated by using a RLS algorithm with a variable forgetting factor. Since the disturbance (when "activated") added a substantial throughput to the system, three b's for varying time-delay and one a were identified on-line. The system was excited by a 16-sample PRBS before the temperature was brought to the nominal operating range of 35°C. Three sets of results using three different tuning configurations were obtained in a single run. The configurations corresponding to the figure number are tabulated below:

Figure	$n1$	$n2$	nu	γ
3.15	1	5	1	.6
3.16	1	5	1	0
3.17	1	10	1	0

Table 3.3: Tuning parameter configurations

In each figure, outlet temperature and setpoint (dotted line) are plotted in the upper portion, steam and cold water (dashed line) flow rates are in the lower portion. Since all three sets of results were obtained in **one** single run and model parameters might drift during different transient responses, the initial setpoint change in each figure is regarded as an excitation for each new set of tuning configuration.

It is apparent from table 3.3 that fig. 3.15 compares the response with γ-weighting to fig. 3.16 which is without γ-weighting. Fig. 3.17 provides a reference for fig. 3.15. Fig. 3.16 has a prediction horizon of only 5, its response is necessarily faster than the other two but the control action is significantly large. Fig. 3.15 shows that

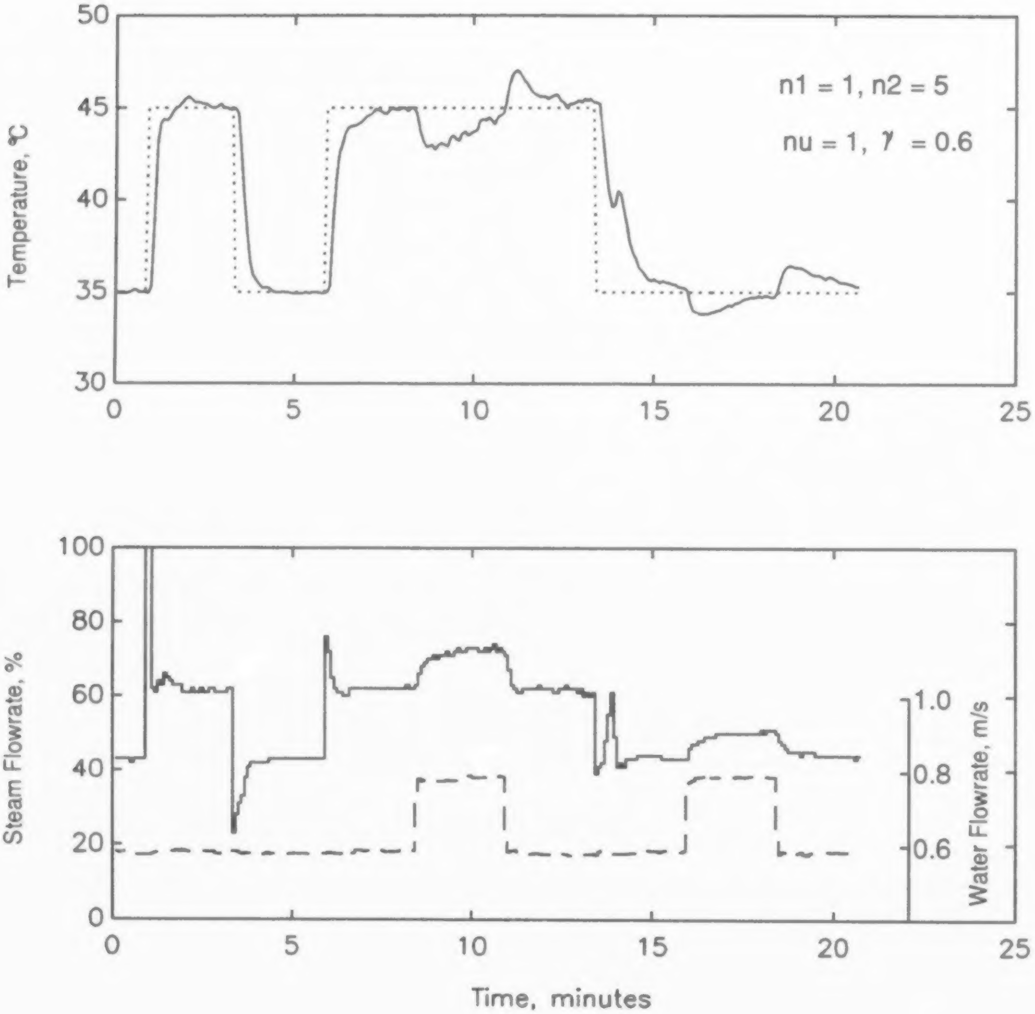

Figure 3.15: Closed-loop performance of continuous stirred-tank heater for ($n1$, $n2$, nu, γ)=(1,5,1,.6)

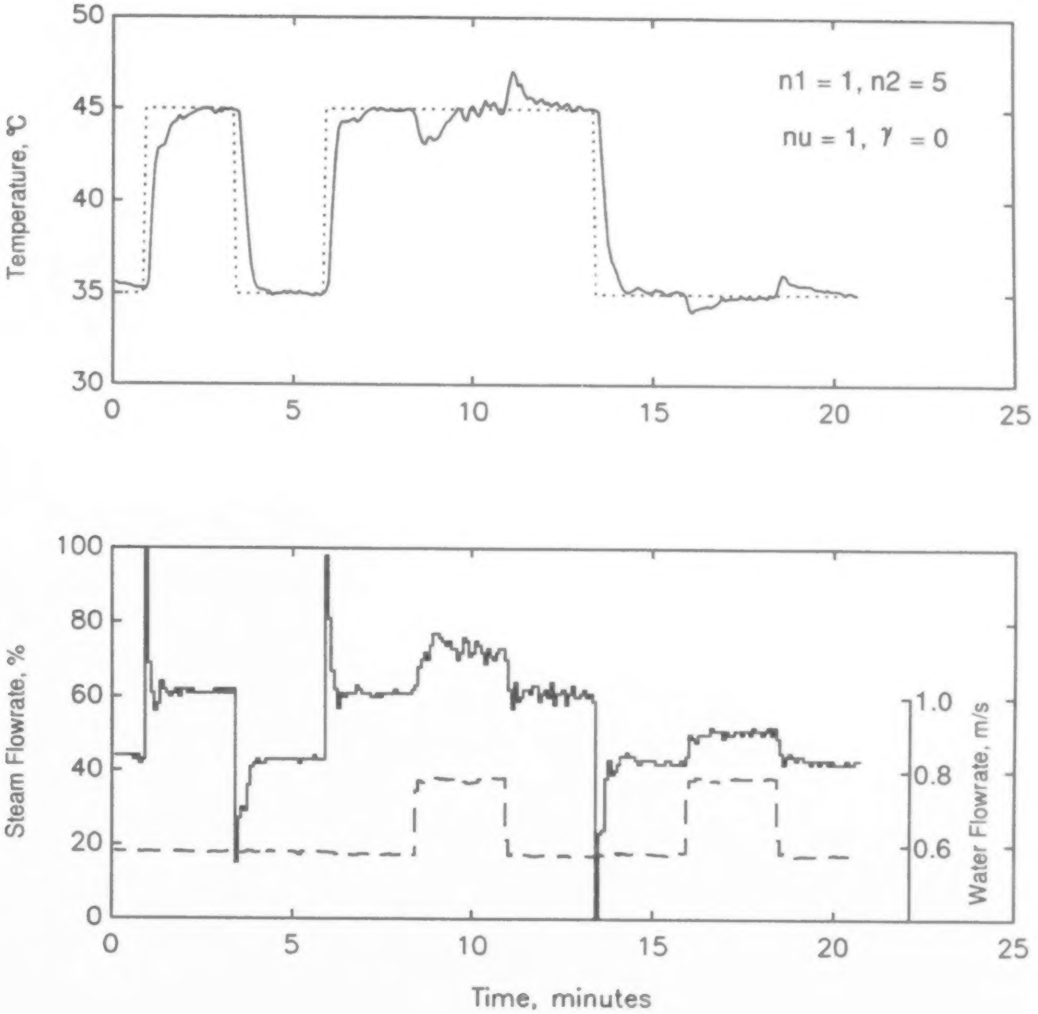

Figure 3.16: Closed-loop performance of continuous stirred-tank heater for ($n1$, $n2$, nu, γ)=(1,5,1,0)

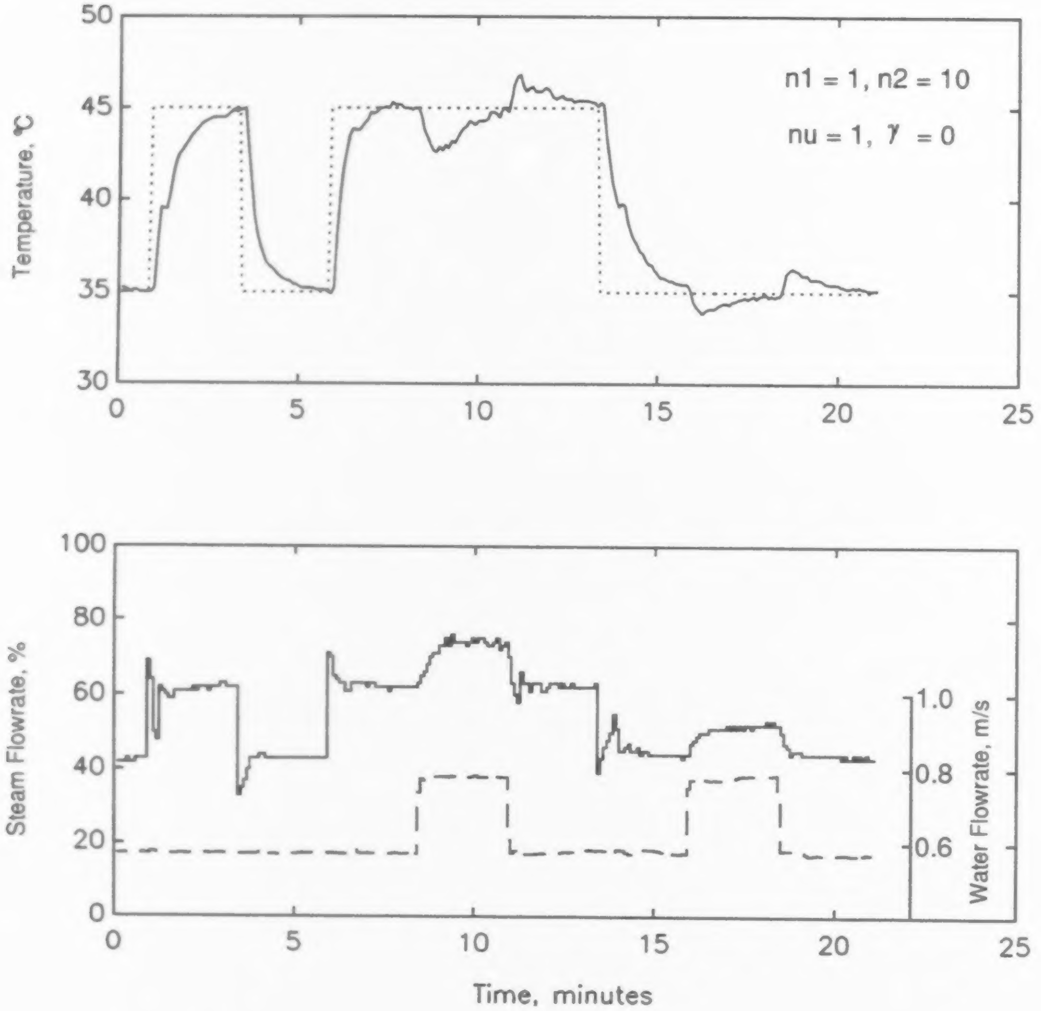

Figure 3.17: Closed-loop performance of continuous stirred-tank heater for ($n1$, $n2$, nu, γ)=(1,10,1,0)

a shorter prediction horizon with γ-weighting provides an effect comparable to the larger prediction horizon in fig. 3.17 both in regulatory and servo control.

3.5 Conclusions

Two tuning strategies are proposed for adjusting the output performance by the steady-state error weighting term while keeping other tuning parameters constant. The first strategy corresponds to a basic dead-beat controller detuned by γ-weighted mean-level controller. The second strategy recommends a large initial value of γ and a short prediction horizon fixed by the rule of discretization. Root-locus results confirm that, in contrast to λ-weighting, γ-weighting does not generate an additional pole in the closed-loop. It also stabilizes open-loop stable non-minimum phase systems. Its equivalence to the output prediction horizon is also shown by both simulations and real experimental evaluations on a continuous-stirred heating system.

Bibliography

Clarke, D.W. and P.J. Gawthrop. Self-tuning control. *IEE Proc.-D*, 126(6):633–640, 1979.

Clarke, D.W., C. Mohtadi, and P.S. Tuffs. Generalised predictive control – Part I. The basic algorithm. *Automatica*, 23(2):137–148, 1987.

Cutler, C.R. and B.L. Ramaker. Dynamic matrix control — a computer control algorithm. In *Proc. JACC*, San Francisco, 1980.

de Keyser, R.M.C. and A.R. van Cauwenberghe. Extended prediction self-adaptive control. In *IFAC Symposium on Identification and System Parameter Estimation*, pages 1255–1260, York, U.K., 1985.

Garcia, C.E. and M. Morari. Internal model control. 1. A unifying review and some new results. *Ind. Eng. Chem. Process Des. Dev.*, 21:308–323, 1982.

Marchetti, J.L., D.A. Mellichamp, and D.E. Seborg. Predictive control based on discrete convolution models. *Ind. Eng. Chem. Process Des. Dev.*, 22:448–495, 1983.

Maurath, P.R., D.A. Mellichamp, and D.E. Seborg. Predictive controller design for SISO systems. In *Proc. American Control Conference*, pages 1546–1552, Paper FP4, 1985.

Maurath, P.R., D.A. Mellichamp, and D.E. Seborg. Predictive controller design for single-input/single-output (SISO) systems. *Ind. Eng. Chem. Process Des. Dev.*, 27:956–963, 1988.

McIntosh, A.R. Performance and tuning of adaptive generalized predictive control. M.Sc. thesis, University of Alberta, 1988.

Mohtadi, C. *Advanced Self-Tuning Algorithms.* D.Phil thesis, University of Oxford, 1987.

Mohtadi, C. and D.W. Clarke. Generalized predictive control, LQ or pole placement: A unified approach. In *Proc. 25th Control and Decision Conference*, pages 1536–1541, Athens, Greece, 1986.

Mosca, E., G. Zappa, and C. Manfredi. Multistep horizon self-tuning controllers: the MUSMAR approach. In *Proc. IFAC 9th World Congress*, pages 935–939, Budapest, 1984.

Peterka, V. Predictor-based self-tuning control. *Automatica*, 20:39–50, 1984.

Qiu, Z., M. Fortuna, and D.G. Fisher. Multi-purpose control system package. Technical report, Dept. of Chem. Eng., University of Alberta, 1988.

Richalet, J., A. Rault, J.L. Testud, and J. Papon. Model predictive heuristic control: applications to industrial processes. *Automatica*, 14:413–428, 1978.

Scattolini, R. and S. Bittanti. On the choice of the horizon in long-range predictive control - Some simple criteria. *Automatica*, 26(5):915–917, 1990.

Shook, D.S., C. Mohtadi, and S.L. Shah. Identification for long range predictive control. *IEE Proc.-D*, 138(1):75–84, 1991.

Xi, Y. Minimal form of a predictive controller based on the step response model. *Int. J. Cont.*, 49:57–64, 1989.

Ydstie, B.E. Extended horizon adaptive control. In *IFAC 9th Triennial World Congress*, pages 911–915, Budapest, Hungary, 1984.

Chapter 4

LRPI with a Terminal Matching Condition[1]

4.1 Introduction

There are two basic ingredients in all long-range predictive control (LRPC) strategies, a specific model structure for finite-horizon output prediction and the minimization of a summation of squared error between multi-step predictions and future setpoints. The major feature distinguishing different strategies lies in the differences in the model structure. The commonly used model structures are either in the form of transfer functions describing the input/output relationships or a model represented by a finite polynomial describing a truncated impulse or step response model. For example, DMC (Cutler and Ramaker, 1980) and MOCCA (Sripada and Fisher, 1985) utilize the step response model whereas MAC (Richalet *et al.*, 1978) employs an impulse response model to formulate the final control law. The difficulty in applying these long-range control laws in adaptive control is that a large number of parameters is usually required to reduce the effect of truncation. Using a standard recursive least squares algorithm for this purpose also implies that a sizable covariance matrix computation is unavoidable.

---

[1] A version of this chapter was presented at the 1991 American Control Conference: Long-range predictive control and identification with steady-state error weighting, Kwok, K.Y. and S.L. Shah, *Proc. American Control Conference*, pp. 2806-2811, 1991.

The success of Generalized Predictive Control (GPC) algorithm (Clarke *et al.*, 1987) in adaptive control stems from the fact that a transfer function model containing significantly fewer parameters is used for future predictions. However, the output predictions based on the diophantine identity in the GPC model structure have a larger variance that grows as the prediction horizon increases (Lu and Fisher, 1990a). The variance makes the long-range projection of the process output less accurate than with other model structures. The inaccuracy is even more severe when a structural mismatch between the model and actual process is present, due to the fact that the least squares approach only aims to provide a minimum variance one-step ahead prediction, *i.e.*

$$J_{LS} = \sum_{t=1}^{T} [\hat{y}(t|t-k) - y(t)]^2 \tag{4.1}$$

Therefore, Shook *et al.* (1991) proposed an adaptive filtering approach to implement a control-relevant long-range predictive identification (LRPI) algorithm based on the dual of the control objective function. Another approach for long-range identification given by Lu and Fisher (1990b) makes use of a non-minimal model structure. However, both approaches are computationally expensive expecially when a large output prediction horizon is utilized.

Since an alternative control objective which reduces the computational load by considering a shorter future projection and the terminal condition has been developed in the preceding chapters, this chapter examines the idea of "dualizing" LRPC with a terminal matching condition to LRPI plus a terminal matching condition. Thus, a "short" output prediction horizon can be used in the mutually-compatible control and identification strategies of LRPC and LRPI so that the resulting adaptive controller is relieved of heavy computations.

4.2 Overall Adaptive Control Objective

4.2.1 Finite Horizon Predictive Control

The overall predictive control objective is to maintain the process output at the desired value by varying a manipulated variable, not only at the current time, but also over a finite-time control horizon. The overall control objective function to be minimized is given as

$$J = \int_0^\infty [y(t) - y_{sp}(t)]^2 dt \tag{4.2}$$

The equivalence to eqn. 4.2 in discrete time is a finite horizon predictive control function,

$$J_{FHPC} = \sum_{j=n1}^{n2} [y(t+j) - y_{sp}(t+j)]^2 \tag{4.3}$$

which aims to keep the future output values at the desired set-point over the time interval from $(t + n1)$ to $(t + n2)$. Eqn. 4.3 can be further decomposed into the following two different objectives plus a cross-product term,

$$J_{FHPC} = J_C + J_{ID} + J_X \tag{4.4}$$

where

$$J_C = \sum_{j=n1}^{n2} [\hat{y}(t+j) - y_{sp}(t+j)]^2 \tag{4.5}$$

$$J_{ID} = \sum_{j=n1}^{n2} [y(t+j) - \hat{y}(t+j)]^2 \tag{4.6}$$

and J_X is the appropriate cross-product term.

Several long-range predictive control laws mentioned in the introduction are successful in meeting part of the objective defined by eqn. 4.4 because they all consider a trajectory of future process outputs rather than just a single or k-step ahead minimum variance control.

To minimize the identification part of the overall objective (eqn. 4.6 for adaptive control, Shook *et al.* (1991) proposed that an appropriate control-relevant identification strategy should provide a model not only for one-step ahead prediction, but over the entire prediction horizon that the control law is required to minimize in the objective function, *i.e.*

$$J_{LRPI} = \frac{1}{T - n2} \sum_{t=1}^{T-n2} \left[\frac{1}{np} \sum_{j=n1}^{n2} (y(t+j) - \hat{y}(t+j \mid t))^2 \right] \tag{4.7}$$

where np is the number of predictions in the horizon and T is the current sampling time. Since optimization based on eqn. 4.7 is non-linear, its implementation in adaptive control application is not practical. Instead, Shook *et al.* (1991) provide a solution for eqn. 4.7 via adaptive filtering, in which a L-filter is used to replace the non-linear optimization. The filter is found through the solution of a spectral factorization technique by equating the power spectrum of eqn. 4.7 to that of the following least squares function,

$$J_{LLS} = \frac{1}{T} \sum_{t=1}^{T} L^2(q^{-1}) \left[y(t+j) - \hat{y}(t+j|t)\right]^2 \tag{4.8}$$

Then the GPC control objective in eqn. 4.5 combined with eqn. 4.8 is able to achieve the overall objective function in eqn. 4.4. However, applying both GPC and LRPI by filtering is computationally expensive for large prediction horizons. On the other hand, a small prediction horizon reduces the computation at the cost of less robustness. In GPC, the process predictions require $n2$ iterations of the diophantine identity. The L-filter is a result of a spectral factorization of a $2 \cdot (n2 - 1)$ polynomial. Large prediction horizons suffer from poor convergence of the spectral factorization computation for the L-filter. Too short a prediction horizon results in a less robust controller and an identification of a wrong-gain model.

4.2.2 Finite Horizon Predictive Control with a Terminal Matching Condition

An alternative to the overall control objective defined by eqn. 4.3 is to minimize over a relatively shorter prediction horizon and, yet at the same time, ensure robustness by introducing an infinite-time term, which is the squared error between the steady-state process output and steady-state setpoint, *i.e.*

$$J_{FHPC,S} = \sum_{j=n1}^{n2} [y(t+j) - y_{sp}(t+j)]^2 + [y(t+\infty) - y_{sp}(t+\infty)]^2 \qquad (4.9)$$

Expanding eqn. 4.9 into different objective functions gives the following expressions,

$$J_{FHPC,S} = J_C + J_{ID} + J_{C,S} + J_{ID,S} + J'_X \qquad (4.10)$$

where

$$\begin{aligned} J_C &= \text{as in eqn. 4.5} \\ J_{ID} &= \text{as in eqn. 4.6} \\ J_{C,S} &= (\hat{y}_s - y_{s,sp})^2 && (4.11) \\ J_{ID,S} &= (y_s - \hat{y}_s)^2 && (4.12) \\ J'_X &= \text{sum of cross-product terms} \end{aligned}$$

Minimization of J_C and $J_{C,S}$ has been achieved by GPC plus steady-state error weighting as described in Chapter 2. J_{ID} is achieved by the non-linear least squares optimization as defined by J_{LRPI}. Heuristically, once the four quadratic terms, J_C, J_{ID}, $J_{C,S}$, and $J_{ID,S}$, are minimized, eqn. 4.10 is also minimized. Therefore, the overall finite horizon predictive control with terminal matching condition criterion can be approximated by:

$$J_{FHPC,S} \approx J_{GPC,S} + J_{LRPI} + J_{ID,S} \qquad (4.13)$$

where

$$\begin{aligned} J_{GPC,S} &= \sum_{j=n1}^{n2} [\hat{y}(t+j) - y_{sp}]^2 + \gamma \sum_{j=1}^{nu} [\hat{y}_s - y_{sp}]^2 && (4.14) \\ J_{LRPI} &= \text{as in eqn. 4.7} \\ J_{ID,S} &= \sum_{t=1}^{T} [y_s - \hat{y}_s|t]^2 && (4.15) \end{aligned}$$

The development of this criterion as presented above is not mathematically rigorous. The cross-product terms are omitted because the problem is otherwise mathematically intractable. (Lu and Fisher (1990) provide justification for the omission of these terms.) Nevertheless, the breakdown in eqn. 4.13 shows that the overall control objective can be achieved by:

- a long-range predictive control law which includes a steady-state error weighting term (eqn. 4.14),
- an identification method which provides a model valid over the entire prediction horizon (eqn. 4.7),
- an additional identification method which supplements the model validity for steady-state value prediction (eqn. 4.15).

Eqn. 4.14 has been explained in detail in Chapters 2 and 3. Shook (1991) has extensively examined eqn. 4.7. The feasibility of realizing eqn. 4.15 is tackled in the next section.

4.3 LRPI with a Terminal Matching Condition

4.3.1 LRPI Algorithm

To derive the L-filter through Parseval's theorem, the process model in eqn. 2.8 is first written in the following j^{th}-step ahead prediction form:

$$\hat{y}(t+j|t) = \frac{F_j}{T}y(t) + \frac{E_j\hat{B}}{T}\Delta u(t+j-k) \tag{4.16}$$

Eqn 4.7 is, then, written in terms of eqn. 4.16:

$$J_{LRPI} = \frac{1}{T-n2}\sum_{t=1}^{T-n2}\left[\frac{1}{np}\sum_{j=n1}^{n2}\left(y(t+j) - \frac{F_j}{T}y(t) - \frac{E_j\hat{B}}{T}\Delta u(t+j-k)\right)^2\right] \tag{4.17}$$

In order to facilitate the analysis of J_{LRPI}, the dependent variable, $y(\cdot)$, is written in terms of the independent variable, $u(\cdot)$ by applying two additional equations to J_{LRPI} as follows.

First, the Diophantine identity from eqn. 2.9 is rearranged as:

$$q^j - \frac{F_j}{T} = q^j\frac{E_j\hat{A}\Delta}{T} \tag{4.18}$$

Second, the true process model for the deterministic case is assumed to be of the form

$$y(t) = \frac{B^0}{A^0}u(t-k) \tag{4.19}$$

The resulting J_{LRPI} in terms of the independent variable, $u(\cdot)$, becomes

$$J_{LRPI} = \frac{1}{T-n2}\sum_{t=1}^{T-n2}\frac{1}{np}\sum_{j=n1}^{n2}\left[\left(\frac{E_j\hat{A}\Delta}{T}\right)\left(\frac{B^0}{A^0} - \frac{\hat{B}}{\hat{A}}\right)u(t+j-k)\right]^2 \tag{4.20}$$

Applying the above procedure to eqn. 4.8 gives

$$J_{LLS} = \frac{1}{T}\sum_{t=1}^{T}\left[\left(\frac{L\hat{A}\Delta}{T}\right)\left(\frac{B^0}{A^0} - \frac{\hat{B}}{\hat{A}}\right)u(t-k)\right]^2 \tag{4.21}$$

Both J_{LRPI} and J_{LLS} are now rearranged in terms of $u(\cdot)$. Since $u(\cdot)$ can be regarded as a signal or time series, the square of this signal in the identification functions is related to the mean power in the signal. Then the power spectrum of eqn. 4.20 is identical to that of eqn. 4.21 as long as the following condition holds:

$$|L(\omega)|^2 = \frac{1}{np} \sum_{j=n1}^{n2} |E_j(\omega)|^2$$

or (using spectral factorization)

$$L(q^{-1})L(q) = \frac{1}{np} \sum_{j=n1}^{n2} E_j(q^{-1})E_j(q) \tag{4.22}$$

Therefore, a standard recursive least squares algorithm combined with the data pre-filter L obtained from eqn. 4.22 is able to accomplish the result as long-range identification (eqn. 4.7). Since the identification of the steady state as shown in eqn. 4.12 is a natural extension from eqn. 4.7, it would be instructive to examine first the properties of L-filter for the purpose of accomplishing the minimization in eqn. 4.12. By examining eqn. 4.22 for $n2$ going to infinity and any arbitrary value of $n1$ which is less than $n2$, the following properties have been observed.

Proposition 1

For all open-loop stable ARMAX models, i.e.

$$\hat{A} \neq 0, \forall\, |q| > 1$$

and

$$n1 = 1 \quad or \quad n1 = n2$$

the model tends from an equation error scheme to an output error one as $n2 \rightarrow \infty$.

Proof

According to Lemma 1 in Chapter 2, polynomial E'_j converges to $\frac{T}{\hat{A}}$ as j tends to infinity. When $n1 = 1$ and $n2$ tends to infinity, L becomes an average of all E'_j. Therefore, the measurement data are filtered by $\frac{L}{T}$ (Δ is absent because of the

ARMAX modeling) and so the combined filter is

$$\begin{aligned}\hat{A}\frac{L}{T}y(t) &= q^{-k}\hat{B}\frac{L}{T}u(t)+\frac{1}{T}d+\xi(t)\\ y(t) &= q^{-k}\frac{\hat{B}}{\hat{A}}u(t)+\frac{1}{T}d+\xi(t)\end{aligned}$$

□□□

Proposition 2

For all open-loop stable ARIMAX models and $n1 = 1$ *or* $n1 = n2$, *the model* **does not** *tend from an equation error scheme to an output error one as* $n2 \to \infty$. *Instead,* L *tends to* E_j *which becomes an infinite stationary polynomial as* $n2$ *tends to infinity.*

Proof

The proof of this proposition is straight forward from Remark 1 in Chapter 2.

□□□

It is inferred from the output error scheme in Proposition 1 that an indirect solution to eqn. 4.12 for the ARMAX model is by having accurate values of $\frac{\hat{B}(1)}{\hat{A}(1)}$ and d. Although Proposition 2 does not render any conclusive remedy to achieve identification at steady state for the ARIMAX model, a comparison between eqns. 4.12 and 4.21 leads to the idea of ensuring the integrity of the process gain, *i.e.*

$$J_{ID,S} = \frac{1}{T}\sum_{t=1}^{T}\left[\left(\frac{B^0(1)}{A^0(1)}-\frac{\hat{B}(1)}{\hat{A}(1)}\right)\Delta u(t-k)\right]^2$$

Thus an accurate estimation of the process gain is an indirect but sufficient method for achieving the identification objective in eqn. 4.12. The next section outlines a simple method for gain estimation in combination with the use of LRPI.

4.3.2 Identification of Process Gain

For on-line estimation of a process gain, any simple stable model can be used because process dynamics are not of major concern. Consider a first order ARMA model as follows,

$$\hat{A}y(t) = b\hat{B}u(t-1) \tag{4.23}$$

where $\hat{A}$ and $\hat{B}$ are identified on-line, and b is a scaling factor such that

$$b\frac{\hat{B}(1)}{\hat{A}(1)} \equiv \frac{B(1)}{A(1)}$$

Dividing $\hat{B}(q^{-1})$ by $\hat{A}(q^{-1})$ generates an impulse response and a residual term,

$$\frac{\hat{B}}{\hat{A}} = (h_0 + h_1 q^{-1} + \cdots + h_{j-1}q^{-j+1}) + \frac{q^{-j}F}{\hat{A}} \tag{4.24}$$

Substituting eqn. 4.24 into eqn. 4.23 for $j = n2$ gives

$$\begin{aligned} y(t) &= b[(h_0 u(t-1) + h_1 u(t-2) + \cdots \\ &\quad + h_{n2-1}u(t-n2) + \frac{F}{\hat{A}}u(t-n2-1)] \\ &= bu^* - ay(t-n2) \end{aligned} \tag{4.25}$$

where u^* is the summation of the impulse response coefficients from 0 to $n2-1$ multiplied by the corresponding control action. Eqn. 4.25 is still in itself a first order representation. A simple recursive least squares algorithm is able to provide the necessary parameters b and a for the overall process gain estimation. The formulation in eqn. 4.25 also coincides with the reduced-order, non-minimal model predictor of Lu and Fisher (1990) in which the impulse response coefficients are user-defined or from previous process identification.

Since the first $n2$ steps of the process are being identified by the LRPI approach, the impulse response coefficients required in the gain estimation model are obtained from the LRPI model.

4.4 Synopsis of an Overall Adaptive Controller

The proposed adaptive controller is a combination of the following three ideas:

- LRPI is accomplished by utilizing the LRPI algorithm by Shook *et al.*(1991),
- the terminal matching condition in identification is indirectly implemented by a gain identification algorithm such as the one outlined in Section 4.3.2,
- LRPC with a terminal matching condition is achieved by GPC with steady-state error weighting developed in Chapter 2.

Fig. 4.1 illustrates the structure of this adaptive controller. The long-range identification portion generates a model for the immediate "long-range" predictions which are used in the gain identification algorithm to estimate the process gain. The control law receives the model and gain estimates, and both are applied to calculate the new corrective action.

4.5 Simulation Studies

In this section, the performance of the controller using the alternative control objective function in eqn. 4.13 (*i.e.* minimizing a shorter output horizon window plus the steady-state error) is illustrated by simulations, and compared with its counterpart in eqn. 4.4 (*i.e.* minimizing a large output horizon window). Rohrs' (1984) third order transfer function,

$$G(s) = \frac{2\text{x}229}{(s+1)(s^2+30s+229)} \tag{4.26}$$

is used to describe the true process. In all simulations, the process is sampled at 10 Hz and estimated by a first order ARMA model.

4.5.1 Open-Loop Identification

In order to evaluate the open-loop identification using LRPI (L-filtering) plus gain estimation approach, a white noise signal with $N(0,1)$ is used to excite the process

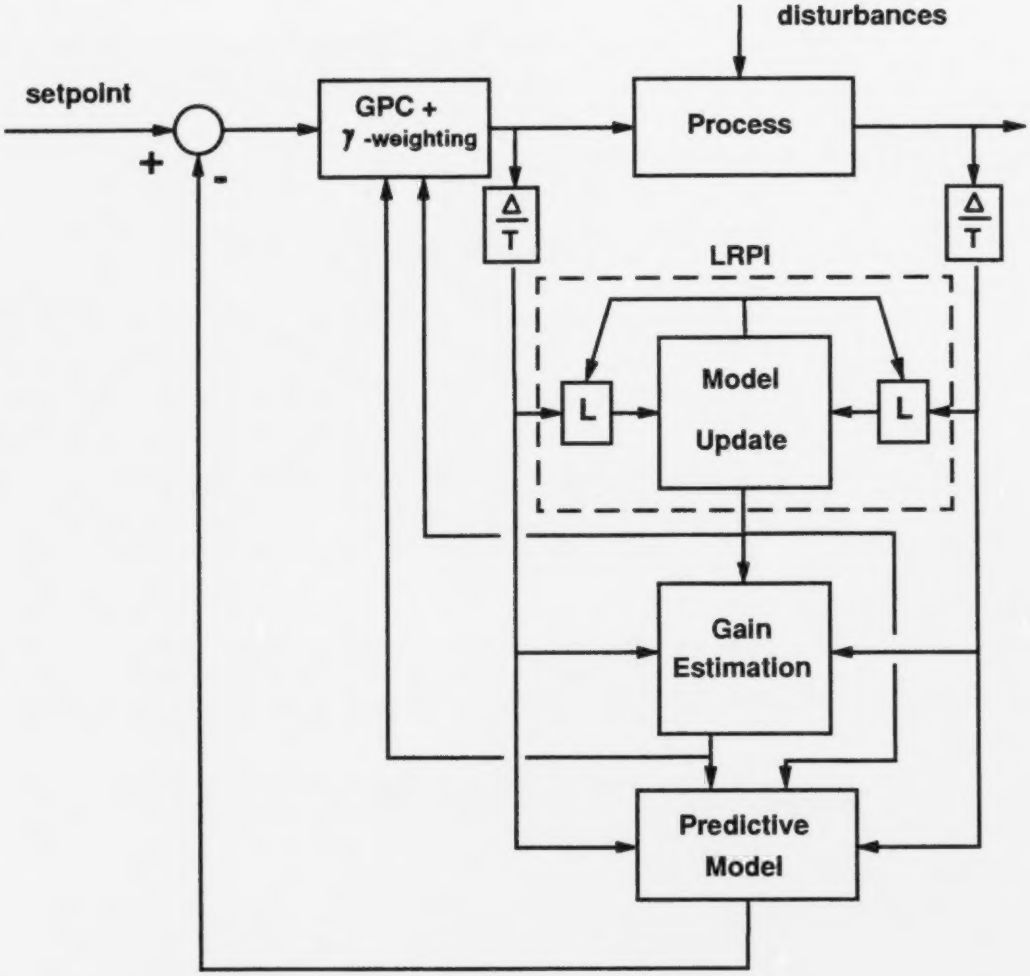
disturbances
setpoint
+
-
GPC +
γ -weighting
Process
Δ
T
Δ
T
LRPI
L
Model
Update
L
Gain
Estimation
Predictive
Model

Figure 4.1: Schematic diagram of an adaptive controller using LRPI/LRPC with terminal matching conditions

for 500 samples. Three different prediction horizons for LRPI are used, 2, 5, and 10. This results in three L-filters separately filtering the input/output data to produce three different first order models. The results are shown in the form of step response trajectories in fig. 4.2. Although $n2 = 2$ produces the closest estimation of process gain (gain=2.11), the poor modeling in the initial step responses disqualifies $n2 = 2$ in closed-loop control (also see next section). The dotted line for $n2 = 5$ gives the greatest deviation (gain=3.84) from the actual process gain of 2, although the initial step response is close to the actual process. When $n2 = 10$ is used, the first order model gives an overall gain of 2.78. As predicted, better identification of a process at low frequencies when significant model-plant mismatch is present in the high frequency region requires a larger $n2$ value which is computationally heavy both in identification and control.

The method outlined in Section 4.3.2 is more successful in estimating the overall gain of the process. Two impulse response polynomials of degree 5 and 10 are used in conjunction with LRPI for $n2 = 5$. These lead to the following structures used in the gain estimation:

$$\begin{aligned} y(t) &= b_5\left[(h_0 u(t-1) + h_1 u(t-2) + \cdots + h_4 u(t-5))\right] + a_5 y(t-5) && (4.27) \\ k_5 &= \frac{b_5(h_0 + \cdots + h_4)}{1 - a_5} && (4.28) \end{aligned}$$

$$\begin{aligned} y(t) &= b_{10}\left[(h_0 u(t-1) + h_1 u(t-2) + \cdots + h_9 u(t-10))\right] + a_{10} y(t-10) && (4.29) \\ k_{10} &= \frac{b_{10}(h_0 + \cdots + h_9)}{1 - a_{10}} && (4.30) \end{aligned}$$

As described previously, both $[b_5, a_5]$ and $[b_{10}, a_{10}]$ are obtained from least squares. The average gains obtained from both eqns. 4.28 and 4.30 are denoted by ○ (averaged gain = 1.68) and + (averaged gain = 1.84) respectively in fig. 4.2. As a result,

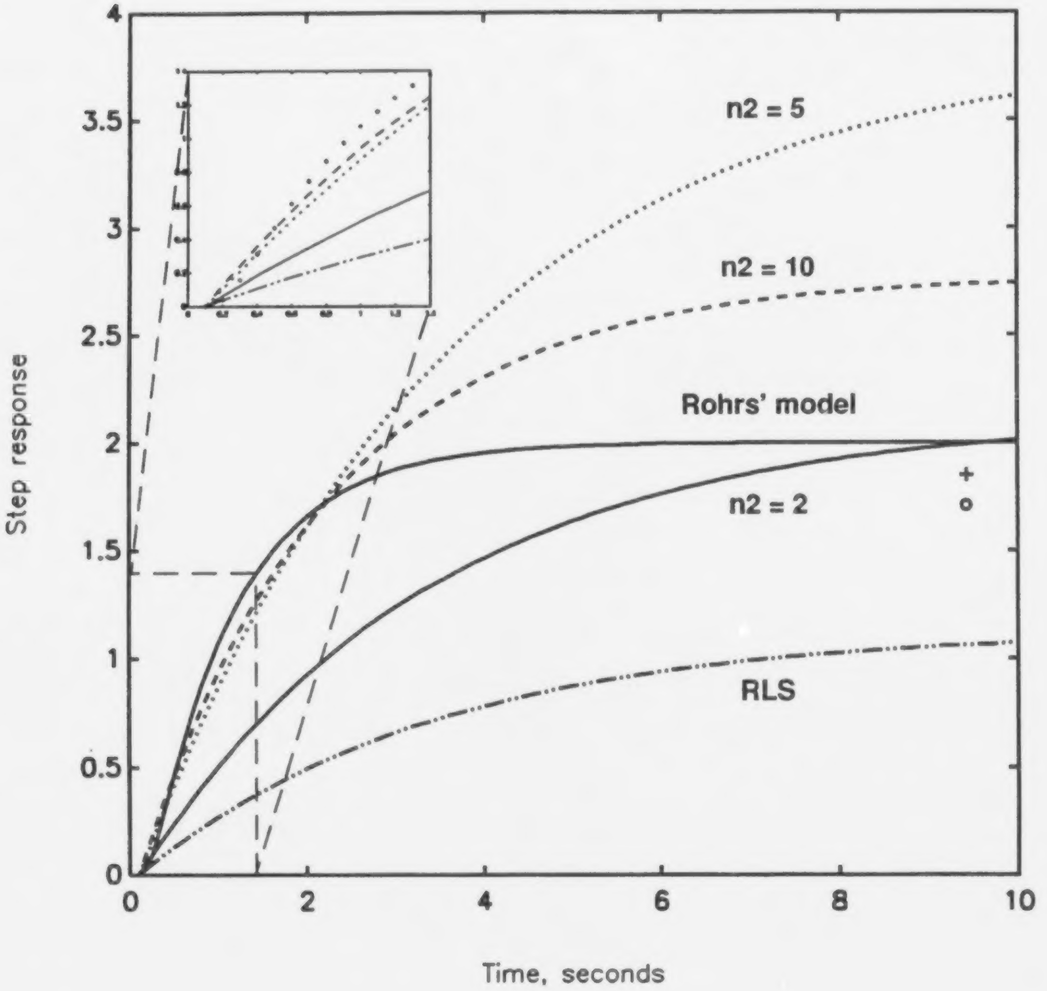

Figure 4.2: Comparison of step response trajectories between Rohrs' model and three LRPI models

the combination of a low-order L-filter and gain estimation will allow a long-range predictive controller to accurately calculate future control moves based on the first few ($n2$) step-responses and the final steady-state value.

4.5.2 Closed-Loop Control

Initial excitation for closed-loop simulations is facilitated by a pseudo-random binary sequence with a magnitude of 1 for 100 samples. When closed-loop control is in effect, the set-point is alternated between 1 and 2 at 50-sample intervals.

Figs. 4.3 and 4.4 respectively show the results of using only GPC and LRPI for the prediction horizons of 2, 5, and 10. In fig. 4.3, poor identification of the first order model and a small prediction horizon for control render the whole controller unstable. When the prediction horizon is changed to 5, the controller stops oscillating. However, overshoot is still present at each set-point change. Better performance is observed in fig. 4.4 when $n2$ is change to 10.

The closed-loop performance obtained using GPC with steady-state error weighting and LRPI with gain estimation is shown in figs. 4.5 and 4.6. With the steady-state error weighting in place, the controller has no difficulty even with a prediction horizon as low as $n2 = 2$. Although oscillations still exist during the initial stage, subsequent control performance is satisfactory. The process response is dampened as γ increases, a trend very similar to that produced by an increase in the prediction horizon.

Fig. 4.6 shows a similar result. Smaller steady-state error weightings are used as the prediction horizon of 5 is set.

The success of the GPC plus steady-state error weighting control in fig. 4.5 requires good estimation of the process gain. The trajectories of gain estimates corresponding to the runs in figs. 4.3 and 4.5 are plotted in figs. 4.7 and 4.8 respectively.

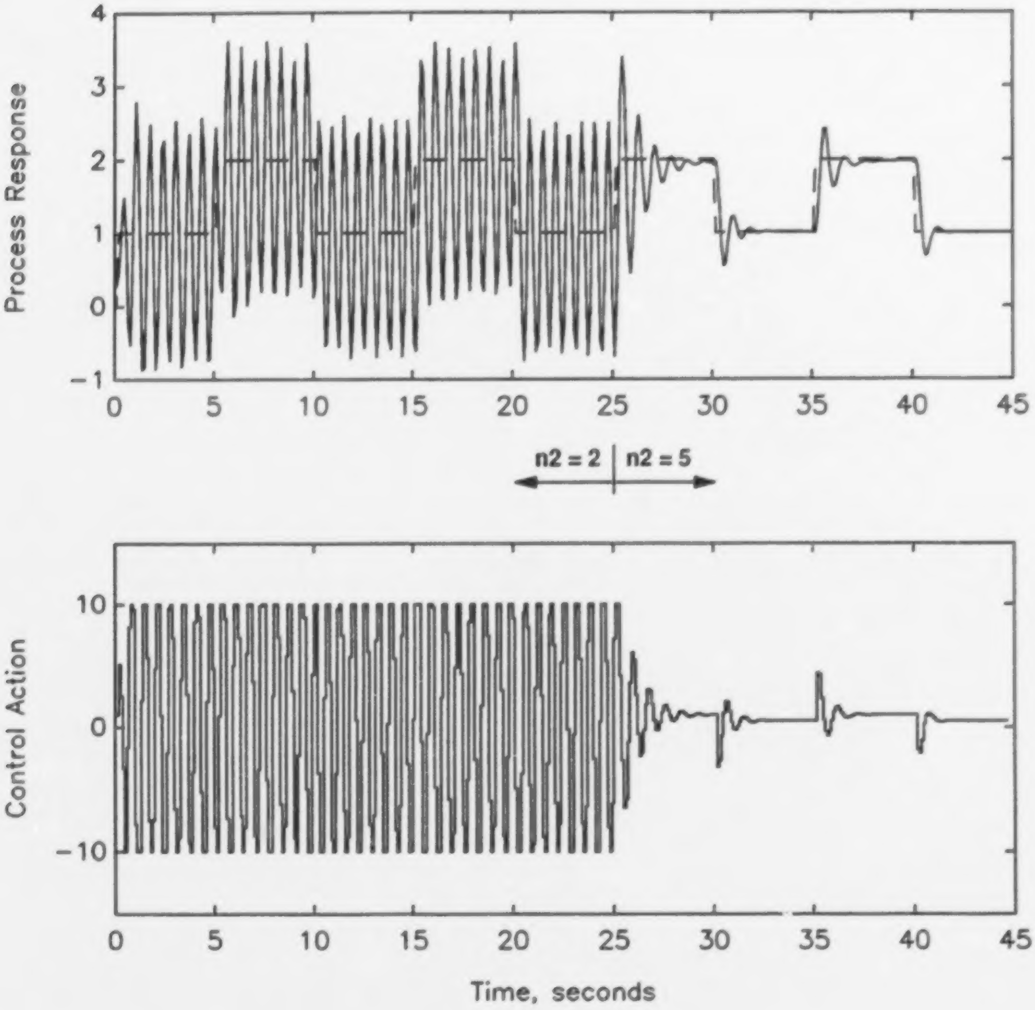

Figure 4.3: Closed-loop response using GPC and LRPI, $n2 = 2$ for $t = 1$ to 250, $n2 = 5$ for $t = 251$ to 450

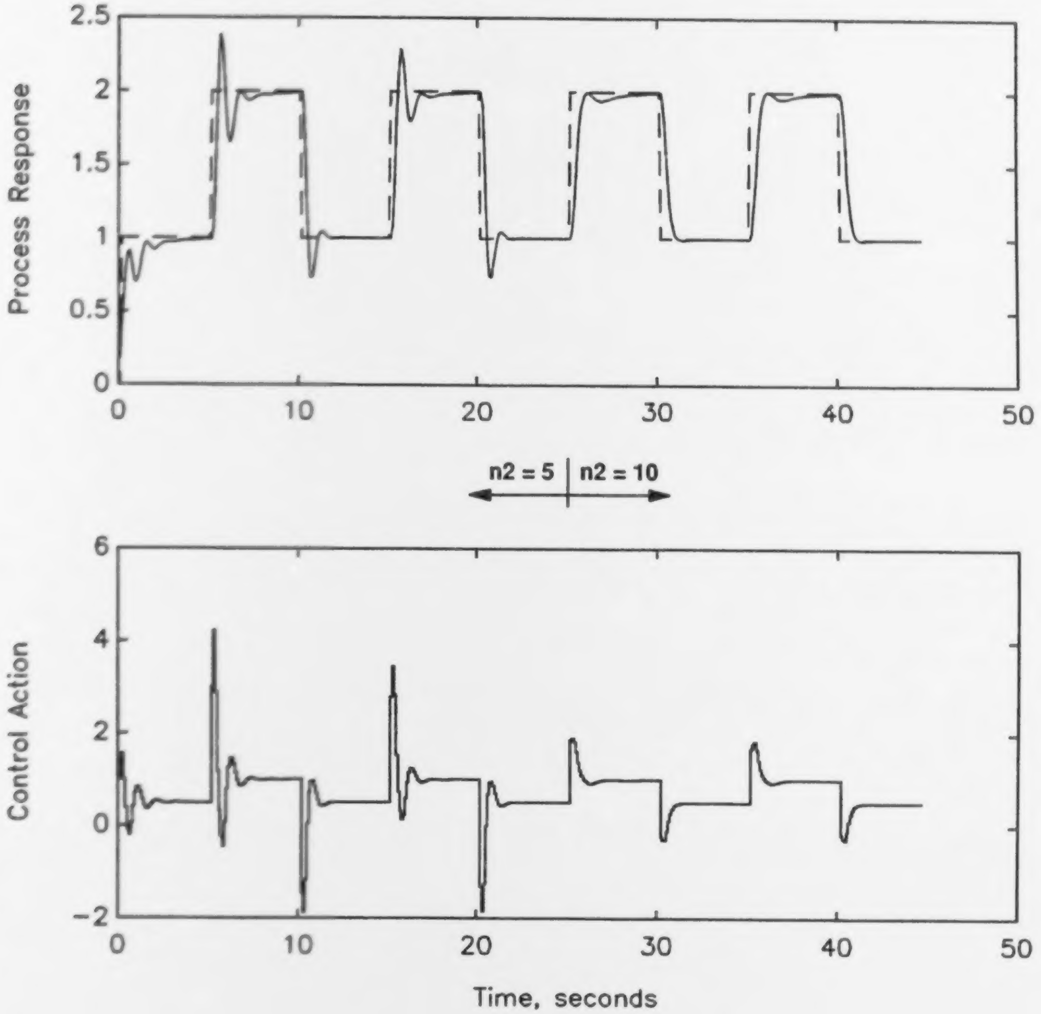

Figure 4.4: Closed-loop response using GPC and LRPI, $n2 = 5$ for $t = 1$ to 250, $n2 = 10$ for $t = 251$ to 450

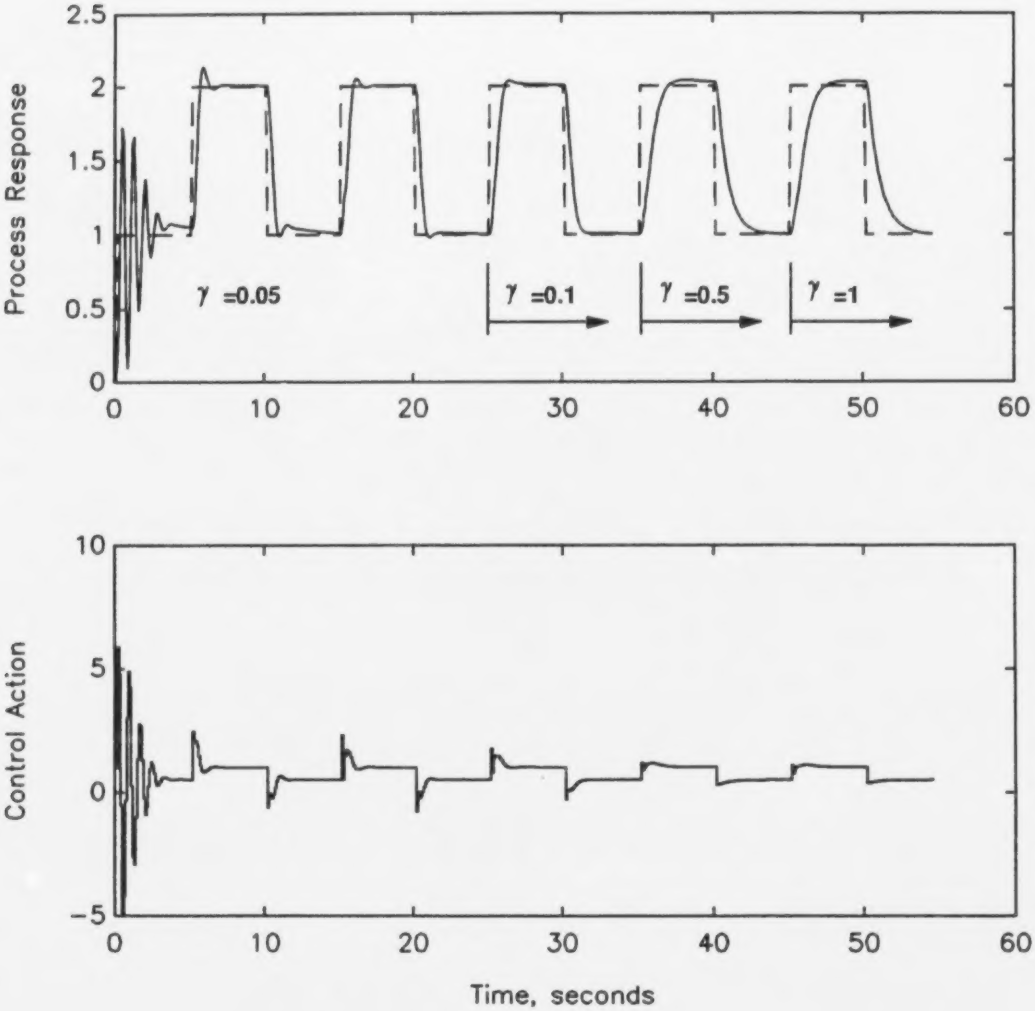

Figure 4.5: Closed-loop response using GPC plus steady-state error weighting and LRPI plus gain estimation, $n2 = 2$, $\gamma = .05, .05, .1, .5, 1$ for each setpoint change

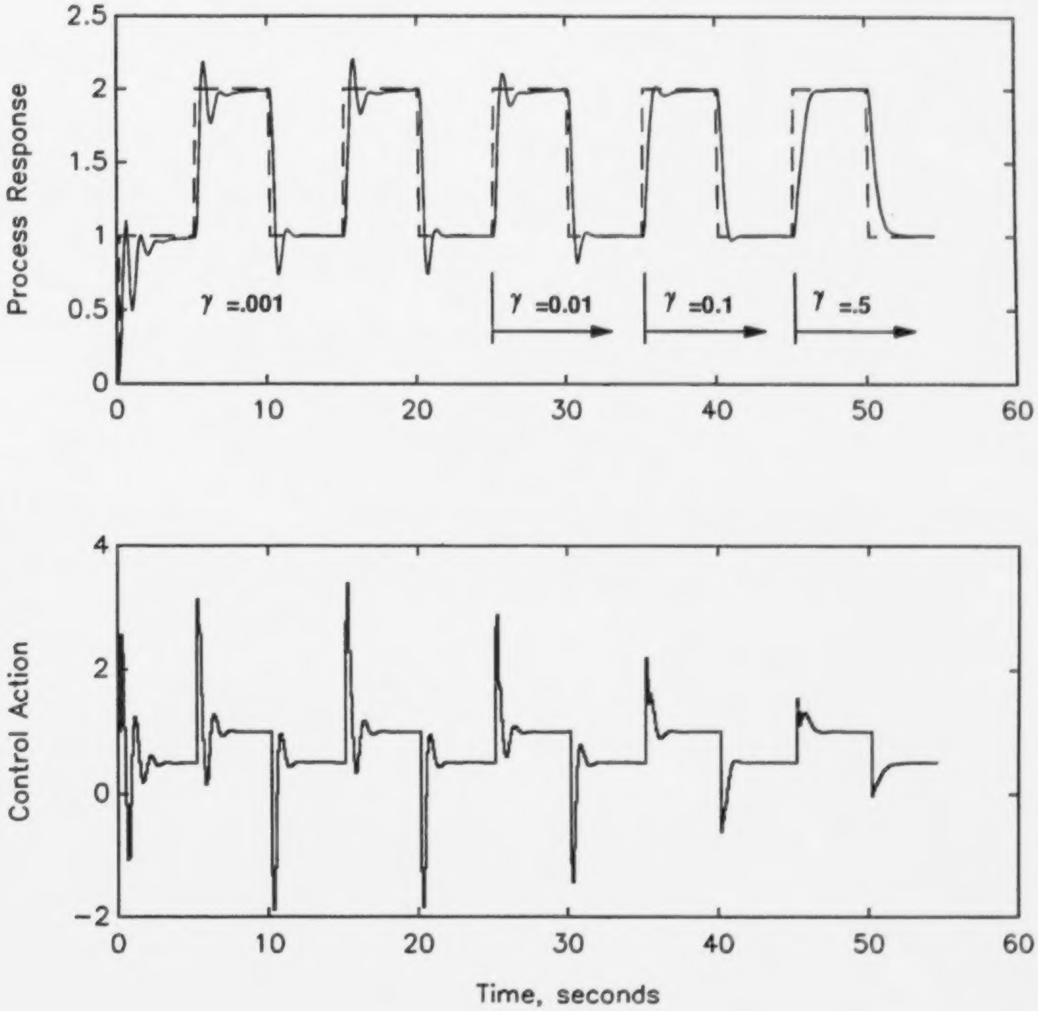

Figure 4.6: Closed-loop response using GPC plus steady-state error weighting and LRPI plus gain estimation, $n2 = 5$, $\gamma = .001, .001, .01, .1, .5$ for each setpoint change

The fast convergence of the estimated gain to the process gain in fig. 4.8 indicates the superior ability of the gain estimation scheme described in Section 4.3.2. The oscillatory gain trajectory shown in fig. 4.7 explains the reason why the control performance in fig. 4.3 is poor. These results lead to the conclusion that accurate gain estimation is critical for any long-range predictive control strategy.

4.6 Conclusions

The overall performance criterion of an adaptive controller is to ensure that the process output tracks the setpoint at all times or at least asymptotically. Minimization of a discrete quadratic cost function over a large finite prediction horizon is computationally "heavy". Conversely, a small prediction horizon reduces the computational load considerably, but results in less robustness. Therefore, an alternative overall performance criterion was proposed to maintain the process output at the setpoint over a relatively shorter prediction horizon and at steady state. The resulting control law is GPC with weighting on the square of steady-state error. The addition of this steady-state error weighting term allows the prediction horizon ($n2$) to be reduced and yet retains robustness as if a large prediction horizon were used. The dual of this control law in identification is a combination of a long-range predictive identification criterion with identification of the steady-state output. An examination of the LRPI algorithm by Shook *et al.* (1991) found that extending the identification horizon to infinity indirectly turns the ARMAX representation of the process from an equation error scheme to an output error scheme. Extension of the identification horizon for an ARIMAX model leads to the idea that identification of the process gain is an indirect method of achieving identification of the steady-state output. Therefore, a simple algorithm for inexpensive on-line gain estimation was developed. The result-

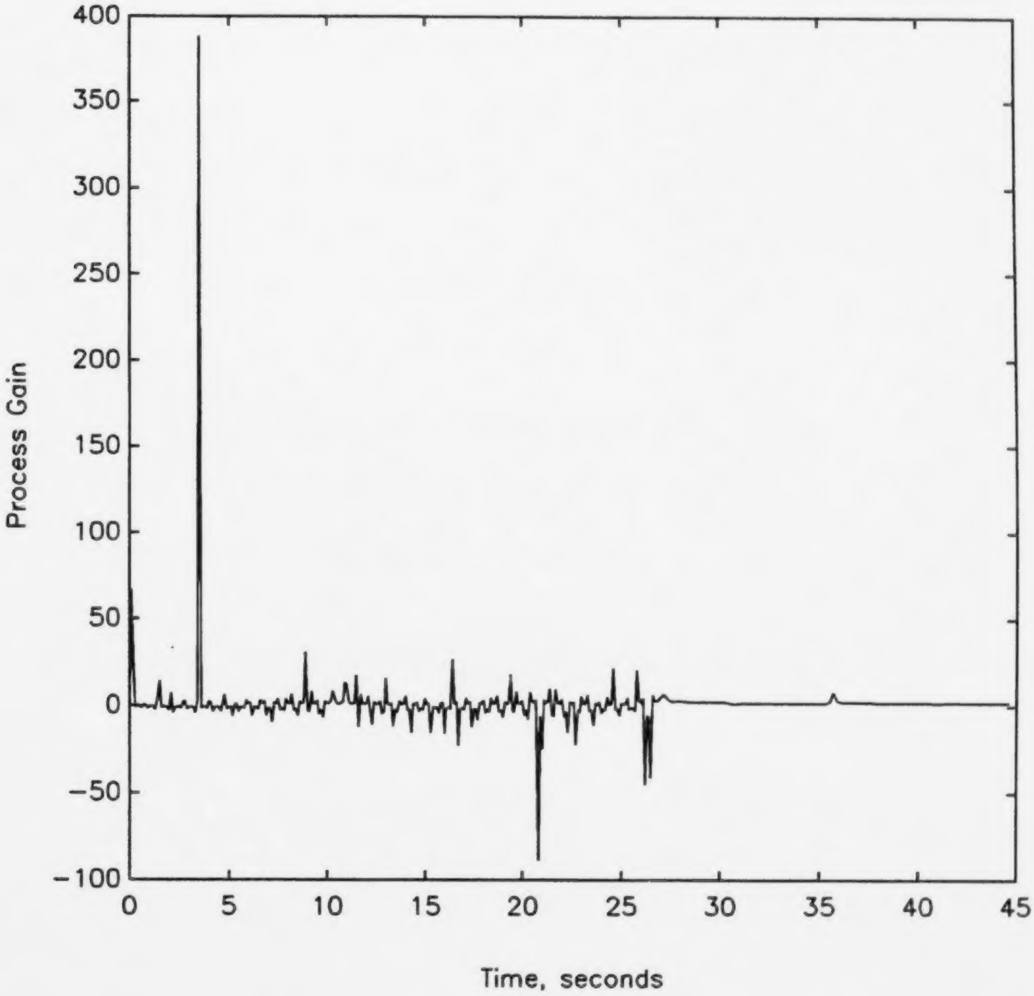

Figure 4.7: Model gain trajectory corresponding to fig. 4.3

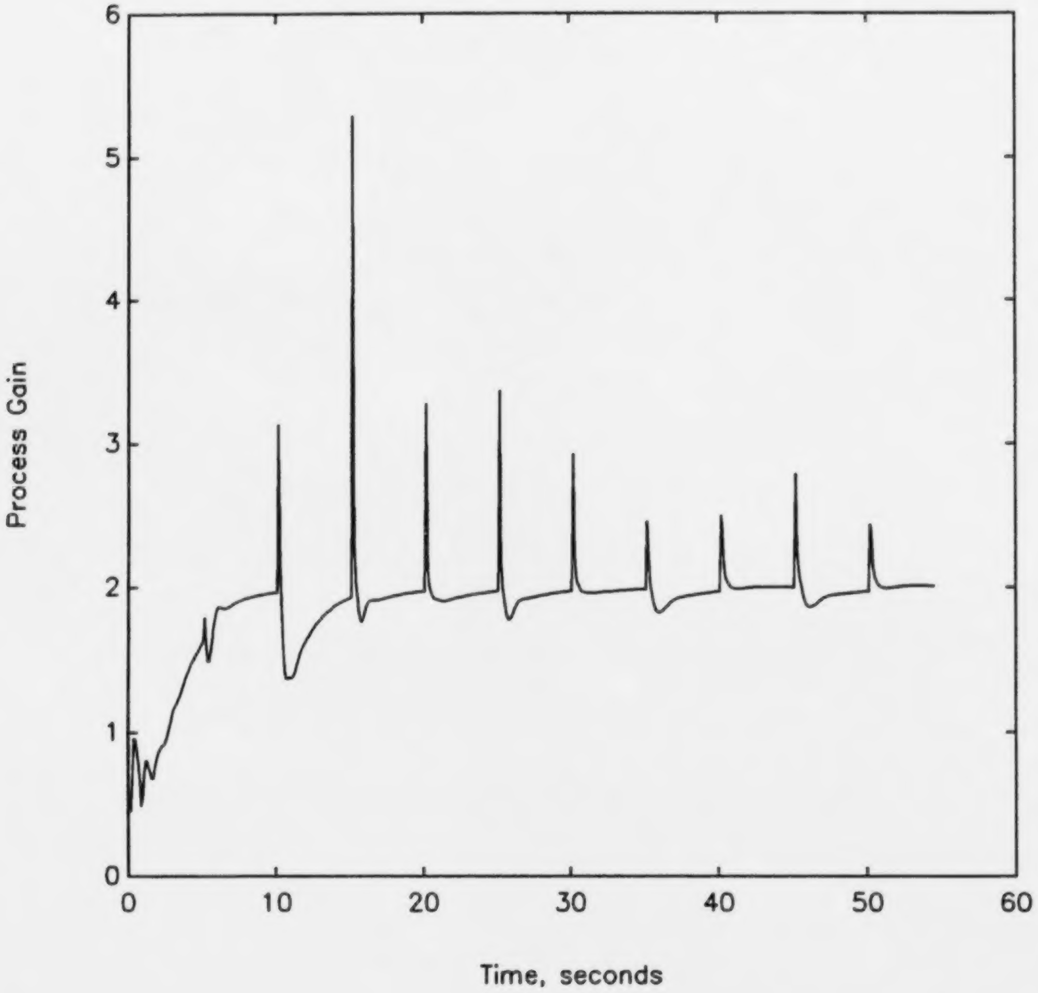

Figure 4.8: Model gain trajectory corresponding to fig. 4.5

ing adaptive controller is a combination of LRPC (e.g. GPC) with steady-state error weighting, the LRPI algorithm of Shook *et al.* (1991), and the process gain estimation algorithm. The performance of this new controller is demonstrated by simulations. The results show that accurate estimation of process gain, combined with steady-state error weighting, improve the overall robustness of the controller even with a very small prediction horizon and large model-plant mismatch.

Bibliography

Clarke, D.W., C. Mohtadi, and P.S. Tuffs. Generalised predictive control – Part I. The basic algorithm. *Automatica*, 23(2):137–148, 1987.

Cutler, C.R. and B.L. Ramaker. Dynamic matrix control — a computer control algorithm. In *Proc. JACC*, San Francisco, 1980.

Lu, W. and D.G. Fisher. Nonminimal model based long range predictive control. In *Proc. American Control Conference*, pages 1607–1613, San Diego, CA, 1990a.

Lu, W. and D.G. Fisher. Nonminimal predictive control. *Chem. Eng. Sci.*, 1990b. In press.

Richalet, J., A. Rault, J.L. Testud, and J. Papon. Model predictive heuristic control: applications to industrial processes. *Automatica*, 14:413–428, 1978.

Rohrs, C.E., M. Athans, L. Valavani, and G. Stein. Some design guidelines of discrete-time adaptive controllers. *Automatica*, 20(5):653–660, 1984.

Shook, D.S. *Identification Issues in Long-Range Predictive Control.* Ph.D. thesis, University of Alberta, 1991.

Shook, D.S., C. Mohtadi, and S.L. Shah. Identification for long range predictive control. *IEE Proc.-D*, 138(1):75–84, 1991.

Sripada, N.R. and D.G. Fisher. Multivariable optimal constrained control algorithm (MOCCA): Part 1. Formulation and Application. In *Proc. International Conference on Industrial Process Modeling and Control*, volume 1, Hangzhou, China, June 1985.

Chapter 5

Automated Blood Pressure Regulation: A Survey

5.1 Introduction

Stable, steady operation of a chemical process is often a non-trivial task. Many times, the process operation is upset by unmeasurable disturbance and interactions from other processes. Varying time-delays and process non-linearities also contribute to the complexity of chemical processes. Early efforts to achieve stable and satisfactory control in the presence of these problems were hindered by the lack of theory and equipment. However, the continuing evolution of computer technology and developments in control theory have made possible the implementation of sophisticated control strategies that can deal with many of these problems.

Since a human body is a dynamic process with many similarities to a chemical process, the concept of using closed-loop control in medicine is not new. Previous practice in performing closed-loop control on physiological variables of a patient was in the form of loose-loop control where medical personnel acted as controllers between the patient (process) and the dosage of an appropriate drug (manipulated variable) (Vozeh and Steimer, 1985). The problems in chemical processes mentioned above also exist in human systems because the human body is subject to numerous deterministic as well as stochastic disturbances. Therefore, manual administration of drugs in intraoperative and postoperative situations to keep important physiological variables within desired limits is difficult and time-consuming. With the progress in automation technology, continuous automated control systems for drug delivery are beginning to appear in clinical medicine. These systems include the control of blood glucose, depth of anesthesia, respiratory variables, intravascular fluids for burn patients, and blood pressure control (Katona, 1982). Since the sensitivity to drug infusion can differ widely for different patients or even changes in the same patient during the course of infusion, an adaptive or self-tuning capability in a control system is beneficial. As a result, many studies in the literature have used advanced adaptive control techniques rather than simple (fixed-parameter) feedback control.

This chapter reviews current developments in the control of blood pressure. Since Sheppard *et al.* (1975) conducted most of the pioneering work in this research area, his work, along with similar developments by others, is described first in the next section. It is followed by a description of the IVAC Titrator[1] which is the first government approved system for automatic nitroprusside delivery. Other contributions are subdivided into different units according to the predictive nature of different control strategies. A survey of Sheppard's work has also appeared in a publication by Linkens (1984).

5.2 Sheppard's System

Sheppard conducted most of the pioneering work in the control of mean arterial pressure using automatic infusion of sodium nitroprusside (SNP). In 1975, a digital system was applied to blood pressure control on thousands of patients (Sheppard *et al.*, 1975). The infusion rate increment was calculated from the proportional and derivative terms of the blood pressure error, modified by a gain scheduling algorithm, incremental limits and decremental terms. All nonlinearities were functions of the error magnitude and sign. They were designed to be aggressive in reducing the infusion rate so as to avoid serious hypotension. However, manual adjustments of the controller gain were often required to improve performance or ensure controller stability (Slate, 1980).

---

[1] IVAC Corporation, San Diego, California

The original system was modified later, based on several modeling studies in the physiological response to SNP infusion. Sheppard and Sayers (1977) assumed a first-order model with a time delay of about one circulation time (20 to 45 seconds). The information was used in the design of a PID controller with a decision table (Sheppard *et al.*, 1979). The clinical evaluation of such a non-linear PID controller showed that further tuning was necessary (Sheppard, 1981). Later modeling investigations (Slate *et al.*, 1979) revealed dynamic properties that could be described as follows:

$$\frac{\Delta P(s)}{I(s)} = \frac{K e^{-\tau_1 s}(1 + \alpha e^{-\tau_2 s})}{(1 + \tau s)} \tag{5.1}$$

where

- ΔP corresponds to the drop of MAP in mmHg induced by SNP infusion rate I in ml hr^{-1},
- K is the sensitivity of the patient to SNP infusion (.25 to 9 mmHg ml^{-1}hr^{-1}),
- τ_1 is the transport delay (20 to 60 seconds),
- τ_2 represents the time delay due to recirculation time (30 to 75 seconds),
- α is the recirculation fraction (0 to 0.4),
- τ is the overall time constant (30 to 60 seconds).

Because of the large range of parameter variations, simulation results using the clinically-used nonlinear PID controller indicated the need for substantial modification and adaptive control (Slate *et al.*, 1980). The design of an improved infusion system including multi-rate filtering of the mean arterial pressure and multi-mode control selection in which parameter adaptation was present was described by Slate and Sheppard (1982a,b). Fig. 5.1 shows the schematic diagram of the system configuration. In this multi-mode control approach, either a linear regulator, nonlinear transient controller, or proportional transient controller was selected to calculate the incremental change every 10 seconds according to the magnitudes of blood pressure deviations and their derivatives. The nonlinear transient controller was a combination of a relay-type controller with a gain scheduler and a Smith-predictor for time delay compensation. The linear regulator was a PD-based controller. The proportional transient controller was used to rapidly turn the controller off if the MAP was far below the desired level. The adaptive feature which is a recursive least squares algorithm modified the gain terms in the nonlinear transient controller and the linear regulator. The gain adaptation routine was executed usually after the initial 15 minutes of control during which other physiological signals were assumed stationary. This controller has been tested on dogs as well as 33 postoperative patients. Besides inaccurate drug gain estimation on several patients due to human operating errors and inadequate initial excitation, this system has shown improved performance compared to the previous design.

The basic control structure of Sheppard's system still belongs to the classical PID feedback concept which requires fine-tuning to achieve desired performance. An example of such a system for MAP regulation is described by Rosenfeldt *et al.* (1986). Their device is a PI based control unit which can be tuned by adjusting three knobs.

Figure 5.1: The schematic diagram of a control system by Slate and Sheppard (1982)

An overview of PID feedback strategy and some tuning methods for implementation in microcomputer based medical devices is given by Westenskow (1986).

de Asla *et al.* (1985) have implemented Sheppard and Slate's control algorithm in a portable unit consisting of a HP-85 microcomputer and an IMED 929 infusion pump. Comparative studies of 49 patients with this system and 37 with experienced intensive-care-unit nurses concluded once again that the computerized drug infusion system is superior to manual control. Reid and Kenny (1987) also applied a different version of Sheppard and Slate's system to control systolic arterial pressure instead of MAP. They found that all patients with computer control had their pressure controlled within ± 10 mmHg of the desired systolic pressure 90 % of the time. Another similar dual-pump control system was applied to deliver SNP and glyceryl trinitrate (Colvin and Kenny, 1989a,b). Because of possible toxic effect of prolonged SNP infusion, trinitrate was used as a primary vasodilator to regulate MAP.

A survey of nurse attitudes towards automatic blood pressure regulation concluded that using computer control allowed more time for nurses to provide other aspects of patient care. However, nurses were critical of the slow response of such a control system and lack of instruction on system operation (Murchie and Kenny, 1988). These results reflect the fact that PID-tuning is necessary and difficult. Especially in the presence of varying time delay and changing dynamics, PID-based controllers must be over-detuned to accommodate uncertainties and increase robustness. This limitation in PID controllers has prompted other attempts in using more advanced control strategies which are described in the sequel.

5.3 IVAC Titrator

A number of investigations have been made in applying predictive and/or adaptive control strategies in the area of biomedical engineering, because the predictive nature is believed to produce faster response and the adaptive nature makes the controller more flexible. Among these investigations, the IVAC Titrator is the first automatic closed-loop MAP control device for SNP delivery which has received approval from the Food and Drug Administration of the United States for postoperative hypotensive therapy. Therefore, it is described in considerable detail.

The project was initiated by the Cleveland Clinic Foundation in Ohio in 1979. The first system was designed to deliver SNP at a rate that was determined by a set of pre-specified range determinations. Details of the range determination were not available. The system was tested on 57 postoperative cardiac patients (Petre *et al.*, 1983). Since the issues of varying time delay and individual patient response were not considered in the controller design, the performance, as indicated by a typical trajectory of MAP, showed large deviations from the desired range. Subsequent development of the system by a joint effort with IVAC Corporation (Meline *et al.*, 1985) added an adaptive minimum variance control strategy (Clarke, 1981). The purpose of this strategy was to find the optimal SNP infusion which minimized the variance of the error between a predicted MAP response and the setpoint:

$$J_{MV} = [\hat{y}(t+k+1) - w(t+k+1)]^2 + \lambda [\Delta u(t)]^2 \tag{5.2}$$

The relationship between MAP and SNP was represented by a time series model in the following form:

$$\begin{aligned} y(t) &= a_1 y(t-1) + a_2 y(t-2) + \cdots + a_{na} y(t-na) + b_{k+1} u(t-k-1) \\ &\quad + b_{k+2} u(t-k-2) + \cdots + b_{k+nb} u(t-k-nb) + \xi(t) \end{aligned} \tag{5.3}$$

where $y(t-i)$ and $u(t-i)$ were MAP response and SNP infusion at time $t-i$, respectively. $\xi(t)$ was a noise term. Predictions of MAP response was made by shifting the time elements in eqn. 5.3 k-units ahead. Adaptation was done by estimating the parameters a_i and b_i using a recursive least squares method.

A comparative study of the performance of the minimum variance control and a particular version of PID control was conducted experimentally (Meline *et al.*, 1986). After the two control strategies were tested with a series of challenges, no significant difference was found between the two on an average basis. However, the minimum variance controller was more aggressive than PID, and its oscillations due to time-delay mismatch were not desirable. Another comparative study between the same PID controller and 10 anesthesiologists was conducted at the University of Utah (Westenskow *et al.*, 1987). A series of challenges by means of other drug infusions to homeostasis was introduced to 10 dogs during induced hypotension by either an anesthesiologist or the automatic controller. The conclusion was that the results using computer control were as good as those obtained by an anesthesiologist who devoted his full attention to blood pressure control. These two comparative studies substantiated the idea that computer control was capable of achieving MAP regulation; and a well-tuned PID was comparable to minimum variance control which was sensitive to modeling error.

The final version of the system was known as IVAC Titrator model 10K which contained two separate control algorithms: the first was a transient control mode, designed to lower MAP to a desired level; the second was a regulatory mode for maintaining MAP at the desired level using the previously-tested PID algorithm. Although details of the transient control mode were proprietary and not available, the control algorithm is believed to be a modified version of the minimum variance control law. A group of 90 postoperative cardiac surgical patients were tested using this system and the results were compared with those from another group of patients. The statistical results showed that about 85 % of MAP regulatory responses with computer control and 61 % of that with manual control were within ±10 % of the desired level (Cosgrove III *et al.*, 1989).

Although this unit has been granted approval for use in postoperative patients, there exists room for improvement such as reducing the variance of MAP during regulatory control, increasing the system's robustness for application in intraoperative patients, or even applying the same MAP control system in other biomedical practices. These incentives have resulted in a lot of efforts in searching for a better MAP control system which is summarized in next section.

5.4 Other Ventures

5.4.1 Single-Step Predictive Optimization

Minimum Variance Based Algorithm

One of the most widely investigated modern control techniques is the minimum variance control strategy (Åström and Wittenmark, 1973; Clarke and Gawthrop, 1975). This type of control strategy, as mentioned in Section 5.3, requires a mathematical model of the effect of SNP on MAP and a quadratic input-output performance index. Its adaptive nature usually appears in the modeling procedure in which the model parameters are frequently adjusted so that optimal performance can be maintained in spite of changes in the process or environment. Because of its predictive nature and simple implementation for adaptation, different versions of the generalized minimum variance control (Clarke and Gawthrop, 1979) applied to MAP control have appeared in the literature. The scope of these studies includes simulations using a blood pressure model (Stern *et al.*, 1981; McInnis and Deng, 1985; Behbehani and Cross, 1991), experiments on dogs (Koivo *et al.*, 1980; Stern *et al.*, 1981; Koivo *et al.*, 1981; Walker *et al.*, 1982; Arnsparger *et al.*, 1983; Meline *et al.*, 1985; Peng *et al.*, 1988), and clinical trials on postoperative patients (Foss *et al.*, 1990) as well as intraoperative patients (Millard *et al.*, 1987; Millard *et al.*, 1988).

A common phenomenon found in most of these results is that SNP infusion appeared to be ringing (or oscillating) and a large MAP variance was observed. This is due to the fact that control performance of any model-based one-step ahead predictive control law is very sensitive to the choice of time delay, k, and model order, na, as in eqn. 5.3. In general, k varies from patient to patient and the true na is infinite. Applying minimum variance control requires *a priori* determination of k and a reduced-order model to approximate physiological dynamics. Poor selections of these values deteriorates control performance and can even lead to instability. Therefore, different configurations of the generalized minimum variance control strategy have been investigated in order to increase stability. The controller by Millard *et al.* (1987) resorted to heavy filtering of both input and output signals. Stern *et al.* (1985) implemented a time-delay estimation algorithm (Kurz and Goedecke, 1981) to identify k on-line, but their study in comparing human performance and the minimum variance control did not show major improvement in using computer control. The control advance moving average controller (CAMAC) which belongs to the class of extended horizon control was implemented to deal with varying and unknown time delay (Voss *et al.*, 1987; Voss *et al.*, 1988). This scheme minimized a similar minimum variance cost function at a specific time advance. The time advance was selected to be either larger than or equal to the actual time delay. However, proper tuning of the time advance was essential, but difficult, especially in the presence of disturbances.

Attempting to obtain a better reduced-order model, Walker *et al.* (1982) compared least squares estimation with extended least squares (ELS) and maximum likelihood (ML) approaches. Although both ELS and ML were capable of identifying the noise model in simulation (Walker *et al.*, 1982), their slow parameter convergence made them undesirable in experimental settings.

The minimum variance type control law is indeed well-known in control engineering. However, its application in a highly non-linear and varying time-delay physiological processes such as MAP control is not appropriate. Because of this rationale, a number of other control techniques have been considered for improved robustness. These include the pole-placement algorithm, multiple-model adaptive control, model reference adaptive control, and long-range predictive control which are described in the followings.

Pole-Placement Algorithm

A pole-placement algorithm functions as a self-configured controller so that the overall closed-loop poles which characterize process dynamics are always placed at prespecified locations. Two versions of this controller have been designed by Mansour and Linkens (1989a,b) using the Slate and Sheppard model. In one development, a Smith predictor was used to eliminate the time delay and reduce model-plant mismatch (Mansour and Linkens, 1989a). Another study made use of an expanded $\hat{B}(q^{-1})$ polynomial to accommodate varying time delay (Mansour and Linkens, 1989b). The model parameters were estimated by a RLS algorithm. Although pole-placement is a more robust control technique than minimum variance control, numerical ill-conditioning due to cancellation of near common factors in an overparameterized model may jeopardize the whole controller performance. Significant model-plant mismatch due to errant estimates of model parameters especially when an expanded $\hat{B}(q^{-1})$ is identified can also render the closed-loop control unstable. Therefore, the authors correctly acknowledged that special care should be taken with respect to parameter "jacketing" when using a self-tuning pole-placement controller.

Multiple-Model Adaptive Control

A multiple-model adaptive controller was designed by He *et al.* (1986) to ensure model integrity. This approach assumed that the plant could be represented by a finite number of models. In He's work, eight models were used to represent MAP response to SNP infusion. The gains of the models were selected to cover the range of MAP sensitivities to SNP infusions given by Slate and Sheppard (1979), whereas all dynamic model parameters were set to their nominal values. A pole-placement based controller associated with each model was designed to satisfy an undershoot less than 10 mmHg and a settling time less than 300 seconds. The resulting closed-loop controller was also subject to a set of non-linear constraints that would limit the dosage being infused into the patient. The final control action was calculated as the weighted sum of all eight outputs from the controllers. As reported by the authors, the largest undershoot and longest settling time in simulation results were about 12 mmHg and 390 seconds, respectively, which violated the authors' own specifications on controller design. When the time delay was underestimated, large deviations from the setpoint were observed.

Martin *et al.* (1987) followed up this design by including a Smith predictor, a low-pass filter, and some minor modifications in their multiple model adaptive controller. The time delay was estimated at every setpoint change ($\geq$ 20 mmHg) by fitting an exponential curve to sequential pressure measurements as the pressure drops by 3 to 10 mmHg. Seven models were assumed instead of eight. The control law was still designed via pole-placement and state-variable feedback. As a result of the time delay estimation, the severe undershoots previously observed in simulation were eliminated. The desired performance characteristics were met in simulation. However, the time-delay estimation function used during setpoint changes could not identify time-delay variations during regulatory control which was considered more common than servo control in MAP regulation.

Although adaptation around a number of fixed models may facilitate the initial identification, the fixed models only account for a limited variation in physiological changes. Moreover, this modeling technique rarely provides an optimal and precise model. The most significant drawback is that multiple-model adaptive controllers are not transferable to control other physiological signals.

Model Reference Adaptive Control

Rather than specifying the performance of a closed-loop controller by placing its poles at certain locations, one could make the performance specification by means of a desired model. This model reference adaptive control has also been considered for MAP control. A state-space model corresponding to a first-order model with a unity gain, time-delay of 30 seconds and 40-second time constant was used as a reference by Sobel *et al.* (1982) and Kaufman *et al.* (1984). The adaptation comes in the controller gain term which would be adjusted at every sampling interval. A Dahlin algorithm was briefly evaluated on dogs by Zhang *et al.* (1988). Pajunen *et al.* (1990) also studied another version of a model reference adaptive controller. The reference model was made adaptive by on-line identification using an ELS algorithm with covariance resetting. The time delay was estimated *a priori* by exciting the process with a pseudo-random binary sequence (PRBS) prior to closed-loop control (Steinmetz, 1987). Their simulation results show that the excitation using PRBS took at least 10 minutes to complete. This long period of probing strategy is usually not practical in clinical situations.

Model reference adaptive control possesses drawbacks similar to pole placement control and falls into the category of single-step predictive control. Also, model reference schemes generally require excessive phase leads which cause difficulties in the presence of noisy signals. This explains why the previous studies mentioned above required such a long period of excitation.

5.4.2 Multi-Step Predictive Optimization

Multi-step predictive optimization strategy inherits the benefit of one-step predictive control and yet is deemed more robust because of its long-range predictive nature. The performance objective considers a trajectory of predicted responses rather than a single point in the future. A generalized optimization index for this class of long-range predictive control is given by the following equation:

$$J_{LRPC} = \sum_{j=n1}^{n2} [\hat{y}(t+j) - w(t+j)]^2 + \sum_{j=1}^{nu} \lambda(j) \, [\Delta u(t+j-1)]^2 \qquad (5.4)$$

where $n1$ and $n2$ represents a horizon of MAP predictions, nu is the number of future SNP infusion changes. A version of this control algorithm is known as generalized predictive control (GPC) (Clarke *et al.*, 1987). With different configurations of the three horizon terms, GPC reduces to many well-known controllers. More importantly, the three simple terms can be tuned to overcome many of the limitations of pole-placement or minimum variance control (McIntosh, 1988). Both Kwok *et al.* (1990, 1991, and see Chapter 6) and Yu *et al.* (1991) applied GPC in the regulation of MAP. The difference between their controller designs was in the model identification. Yu *et* *al.* included GPC in their multiple-model adaptive controller for control of MAP and CO. In contrast, the present study (Chapter 6) considered a truly adaptive model-based predictive MAP controller. An advanced control-relevant identification method for long-range predictive control (Shook *et al.*, 1991) was also incorporated by Kwok *et al.* in a later study which is discussed in detail in Chapter 7.

5.4.3 Multiple Drug Infusion

The MAP is only an indication of the vascular resistance. The main control objective is to provide and maintain the blood flow required by different parts of the body after undergoing cardiac surgery. Therefore, several attempts were initiated to simultaneously control the MAP and other important physiological parameters. Of most interest is the control of cardiac output (CO) because combining MAP and CO gives rise to a quantitative value of the systemic vascular resistance. This problem was well explained from the perspective of an anesthesiologist and a control engineer by Roy (1982).

Cardiac output measures the overall blood flow in the circulatory system, which is the amount of blood pumped by each ventricle of the heart in a unit period of time. The current clinical method of measuring CO is by means of thermodilution performed intermittently. Because of a lack of continuous CO measurement, Serna *et al.* (1983) first investigated this control problem by treating MAP/SNP and CO/dopamine (DOP) as two SISO control loops. Since then, several advanced control strategies have been considered for a true multivariable drug infusion system. These include: adaptive pole-placement, model reference adaptive control and Richalet's model algorithmic controller (1978) in a comparative study via simulation (Lau *et al.*, 1984), as well as multivariable generalized minimum variance control on an electrical analog model of the circulatory system (McInnis and Deng, 1985). However, the results were not conclusive for two reasons. First, the performance of different control strategies largely depends on how the strategy is tuned. Second, the multivariable model used for the simulation study was too simple to reflect the complex drug dynamics. Voss *et al.* (1986) also presented a multivariable version of CAMAC and briefly applied it to animal studies. But their controller required 13 to 30 minutes for initial identification and a sampling time of 40 or 60 seconds (Voss *et al.*, 1988), both of which were inappropriate for clinical applications.

Recently Mansour and Linkens (1990) considered two nonpulsatile models of the cardiovascular systems for the design of a two-input, two-output, multivariable, self-tuning controller. Two pole-placement algorithms, an explicit pole-placement algorithm by Prager and Wellstead (1981) and a pole-zero placement algorithm by Sirisena and Teng (1986), were tested for the control of systemic resistance and cardiac output by manipulating SNP and DOP infusions. The interacting dynamics between the drug infusions (i.e. SNP and DOP) and the physiological outputs (i.e. systemic resistance and cardiac output) were represented by the cardiovascular models given by Moller *et al.* (1983) and Wesseling *et al.* (1982). In their closed-loop simulation studies, application of a multivariable self-tuning controller was demonstrated to be feasible in a complex human cardiovascular system. However, as the authors have stated, their studies "have used fixed parameters representing a typical subject". The controller, therefore, needs to be evaluated under interpatient and intrapatient parameter uncertainty.

Without any breakthrough in reliable frequent CO measurement, the studies in a multivariable MAP and CO controller tend to focus more on the demonstration of the control algorithm rather than the actual design of a clinically acceptable device. Since CO is available at a much slower rate than MAP, inferential control may be able to handle this multi-rate control problem. In this case, MAP and another CO-related physiological signals are treated as primary variables for control purpose. CO is then measured at a slower rate and used to correct any modeling error.

5.4.4 Expert and Fuzzy Control System

Expert control offers a new direction in solving process control engineering problems which normally rely on numerical computations. Based on the fuzzy set theory, the concept of an expert system is to turn human experience and knowledge into a group of fuzzy control rules. These rules form a knowledge base and are applied to the input signals for making a new decision. Therefore, the input signals must first be fuzzified before being processed by the knowledge base. The decision is, then, defuzzified into a value which is usually an incremental change of the manipulated variable. A simple example of fuzzy rules for MAP control was a look-up table in which the infusion rate incremental steps were determined by what range the MAP deviation fell into (Sheppard, 1980; Packer *et al.*, 1987). Other full scale fuzzy controllers were described by Ying *et al.* (1988) and Yamashita *et al.* (1988). An intelligent alarm knowledge base was also developed by Fukui and Masuzawa (1989) to differentiate false MAP readings from true ones.

Expert control systems have been widely considered in clinical practices because they tend to match the judgment of medical personnel (Linkens and Hacisalihzade, 1990). However, their performance may not outperform that of the advanced control strategies described in Section 5.4. On the other hand, many advanced controllers require a supervisory system to make decisions on model identification and minor adjustments in tuning parameters especially for adaptive type systems. It would be appealing to combine the strong points of both expert systems and advanced control strategies together so that the former deals with the overall control system and lets the latter perform tight regulation.

5.5 Conclusions

Multi-disciplinary efforts based on control engineering and physiology have been merged to generate a number of successful applications in the practice of medicine. Mean arterial pressure control is one of the active areas which has received much attention both from the academia and practitioners. Since Sheppard's pioneering work in modeling the dynamic response of mean arterial pressure to vasoactive drug infusions, several versions of blood pressure controllers based on PID control have been reported and applied in postoperative hypotensive therapy. One of these systems, known as the IVAC Titrator, eventually received approval from the U.S. Food and Drug Administration for postoperative therapy. However, proper tuning of PID controllers especially in the presence of varying drug response time and changing body dynamics is still a hindrance to broader application in clinical settings. Often times, PID controllers are over-detuned to accommodate uncertainties and increase robustness. As a result, other advanced control strategies have been considered for mean arterial control.

Because of complex physiological dynamics, a lot of research activity has focused on adaptive and predictive control strategies. However, single point predictive controllers with an adaptive mechanism using recursive least squares parameter estimation often resulted in oscillatory responses. Move suppression techniques such as heavy filtering and control weighting were necessary to reduce overly vigorous control action. An alternative to these single point predictive control strategies is multi-step predictive control and identification schemes. Since multi-step optimization considers a trajectory of mean arterial pressure predictions, the resulting controller performance is more robust for time-varying dynamic systems including varying time-delays.

Multiple drug infusion is indeed a logical step towards full automation of drug therapy. But existing limitations in bio-sensors such as measuring continuous cardiac output have hindered further implementation of a lot of valuable research results from the control area.

Knowledge-based fuzzy controllers represent a new approach to many control areas where parametric models are not available for calculating the manipulated variable. In blood pressure control, an expert system could look after the initial control period while identification is underway, supervise the overall controller when artifacts are detected, and perform minor tuning or override control when extraordinary dynamics are encountered. Therefore, using an advanced control strategy such as multi-step predictive control to provide tight control while the overall system is supervised by an expert system is very desirable not only in biomedical control but also in other industrial applications.

Bibliography

Arnsparger, J.M., B.C. McInnis, J.R. Glover, and N.A. Normann. Adaptive control of blood pressure. *IEEE Trans. Biomed. Eng.*, BME-30(3):168–176, March 1983.

Behbehani, K. and R.R. Cross. A controller for regulation of mean arterial blood pressure using optimum nitroprusside infusion rate. *IEEE Trans. Bio. Eng.*, 38(6):513–521, June 1991.

Clarke, D.W. Implementation of self-tuning controllers. In C.J. Harris and S.A. Billings, editors, *Self-tuning and adaptive control:theory and applications*, pages 144–165. IEEE Control Engineers, London, 1981.

Clarke, D.W. and P.J. Gawthrop. Self-tuning controller. *IEE Proc.-D*, 122(9):929–934, 1975.

Clarke, D.W. and P.J. Gawthrop. Self-tuning control. *IEE Proc.-D*, 126(6):633–640, 1979.

Clarke, D.W., C. Mohtadi, and P.S. Tuffs. Generalised predictive control – Part I. The basic algorithm. *Automatica*, 23(2):137–148, 1987.

Colvin, J.R. and G.N.C. Kenny. Automatic control of arterial pressure after cardiac surgery. *Anaesthesia*, 44:37–41, 1989a.

Colvin, J.R. and G.N.C. Kenny. Development and evaluation of a dual-pump microcomputer-based closed-loop arterial pressure control system. *Int. J. Clin. Monit. Comput.*, 6(1):31–35, January 1989b.

Cosgrove III, D.M., J.H. Petre, J.L. Waller, J.V. Roth, C. Shepherd, and L.H. Cohn. Automated control of postoperative hypertension: A prospective, randomized multicenter study. *Ann. Thorac. Surg.*, 47:678–683, 1989.

de Asla, R.A., A.M. Benis, R.A. Jurado, and R.S. Litwak. Management of postcardiotomy hypertension by microcomputer-controlled administration of sodium nitroprusside. *J. Thorac. Cardiovasc. Surg.*, 89:115–120, 1985.

Foss, B.A., H. Hogner, and S. Rising. Adaptive blood pressure control. In *Proc. 11th IFAC World Congress*, volume 4, pages 7–11, Estonia, USSR, August 13–17 1990.

Fukui, Y. and T. Masuzawa. Knowledge-based approach to intelligent alarms. *J. Clin. Monit.*, 5:211–216, 1989.

He, W.G., H. Kaufman, and R. Roy. Multiple model adaptive control procedure for blood pressure control. *IEEE Trans. Biomed. Eng.*, BME-33(1):10–19, January 1986.

Katona, P.G. Automated control of physiological variables and clinical therapy. *CRC Critical Reviews in Biomedical Engineering*, 8(4):281–310, 1982.

Kaufman, H., R. Roy, and X. Xu. Model reference adaptive control of drug infusion rate. *Automatica*, 20(2):205–209, 1984.

Åström , K.J. and B. Wittenmark. On self tuning regulators. *Automatica*, 9:185–199, 1973.

Koivo, A.J., D. Larnard, and R. Gray. Automated blood pressure control in dogs using a microprocessor. In *Proc. IEEE International Symposium on Circuits and Systems*, volume 2, pages 474–477, Houston, Texas, 1980.

Koivo, A.J., D. Larnard, and R. Gray. Digital control of mean arterial blood pressure in dogs by injecting a vasodilator drug. *Annals of Biomedical Engineering*, 9:185–197, 1981.

Kurz, H. and W. Goedecke. Digital parameter-adaptive control of processes with unknown dead time. *Automatica*, 17(1):245–252, 1981.

Kwok, K.Y., A.S. Clanachan, B. Finegan, and S.L. Shah. Long range predictive control of arterial blood pressure. In *Proc. 29th IEEE Conference on Decision and Control*, pages 2800–2805, Honolulu, Hawaii, December 5-7 1990.

Kwok, K.Y., R.K. Mutha, S.L. Shah, A.S. Clanachan, and B.A. Finegan. Constrained long-range adaptive predictive control of arterial blood pressure. *International Journal of Adaptive Control and Signal Processing*, 5(6):363–374, 1991.

Lau, K., H. Kaufman, V. Serna, and R. Roy. Evaluation of three adaptive control procedures for multiple drug infusion. In *Proc. of the 23rd Conference on Decision and Control*, pages 392–393, 1984.

Linkens, D.A. Computer control in biomedicine. In S. Bennett and D.A. Linkens, editors, *Real-Time Computer Control*, chapter 14. P. Peregrinis, 1984.

Linkens, D.A. and S.S. Hacisalihzade. Computer control systems and pharmalogical drug administration: a survey. *J. Med. Eng. & Tech.*, 14(2):41–54, 1990.

Mansour, N.E. and D.A. Linkens. A self-tuning Smith predictor applied to the control of blood pressure: simulation studies. *Computer Methods and Programs in Biomedicine*, 29:21–29, 1989a.

Mansour, N.E. and D.A. Linkens. Pole-assignment self-tuning control of blood pressure in postoperative patients: a simulation study. *IEE Proc.-D*, 136(1):1–11, January 1989b.

Mansour, N.E. and D.A. Linkens. Self-tuning pole-placement multivariable control of blood pressure for post-operative patients: a model-based study. *IEE Proc.-D*, 137(1):13–29, January 1990.

Martin, J.F., A.M. Schneider, and N.TY Smith. Multiple-model adaptive control of blood pressure using sodium nitroprusside. *IEEE Trans. Biomed. Eng.*, BME-34(8):603–611, August 1987.

McInnis, B.C. and L.Z. Deng. Automatic control of blood pressures with multiple drug inputs. *Annals of Biomedical Engineering*, 13:217–225, 1985.

McIntosh, A.R. Performance and tuning of adaptive generalized predictive control. M.Sc. thesis, University of Alberta, 1988.

Meline, L.J., D.R. Westenskow, N.L. Pace, and M.N. Bodily. Computer-controlled regulation of sodium nitroprusside infusion. *Anesth. Analg.*, 64:38–42, 1985.

Meline, L.J., D.R. Westenskow, A. Somerville, R.T. Wernick, J. Jacobs, and N.L. Pace. Evaluation of two adaptive sodium nitroprusside control algorithms. *J. Clin. Monit.*, 2:79–86, 1986.

Millard, R.K., P. Hutton, E. Pereira, and C. Prys-roberts. On using a self-tuning controller for blood pressure regulation during surgery in man. *Comput. Biol. Med.*, 17(1):1–18, 1987.

Millard, R.K., C.R. Monk, and C. Prys-Roberts. Surgery using a volatile anaesthetic. *IEE Proc.-D*, 135(2):95–105, March 1988.

Moller, D., D. Popovic, and G. Thiele. Modelling, simulation and parameter-estimation of the human cardiovascular system. In Irmfried and Hartmann, editors, *Advances in control systems and signal processing*, volume 4, 1983.

Murchie, C.J. and G.N.C. Kenny. Nurse attitudes to automatic computer control of arterial pressure. *Intensive Care Nursing*, 4:112–117, 1988.

Packer, J.S., D.G. Mason, J.F. Cade, and S.M. McKinley. An adaptive controller for closed-loop management of blood pressure in seriously ill patients. *IEEE*

Trans. Biomed. Eng., BME-34(8):612–616, August 1987.

Pajunen, G.A., M. Steinmetz, and R. Shankar. Model reference adaptive control with constraints for postoperative blood pressure management. *IEEE Trans. Bio. Eng.*, 37(7):679–687, July 1990.

Peng, L., G. Wang, S. Zhang, and H. Hou. An adaptive controller for blood pressure. In *8th IFAC/IFORS symp. on Ident. syst. para. est.*, volume 3, pages 1835–1839, Beijing, China, 1988.

Petre, J.H., D.M. Cosgrove, and F.G. Estafanous. Closed loop computerized control of sodium nitroprusside. *Trans. Am. Soc. Artif. Intern. Organs.*, 29:501–505, 1983.

Prager, D.L. and P.F. Wellstead. Multivariable pole-placement self-tuning controller. *IEE Proc.-D*, 128(1):9–18, 1981.

Reid, J.A. and G.N. Kenny. Evaluation of closed-loop control of arterial pressure after cardiopulmonary bypass. *Br. J. Anaesth.*, 59:247–255, 1987.

Richalet, J., A. Rault, J.L. Testud, and J. Papon. Model predictive heuristic control: applications to industrial processes. *Automatica*, 14:413–428, 1978.

Rosenfeldt, F.L., V. Chang, M. Grigg, S. Parker, R. Cearns, M. Rabinov, and W.G. Xu. A closed loop microprocessor controller for treatment of hypertension after cardiac surgery. *Anaesth. Intens. Care*, 14:158–162, 1986.

Roy, R.J. Adaptive cardiovascular control using multiple drug infusions. In *Proc. IEEE Engineering in Medicine and Biology Society*, pages 459–464, 1982.

Serna, V., R. Roy, and H. Kaufman. Adaptive control of multiple drug infusions. In *Proc. American Control Conference*, pages 22–26, San Francisco, CA, 1983.

Sheppard, L.C. Computer control of the infusion of vasoactive drugs. In *IEEE 1980 International Symposium on Circuits and System*, pages 469–473, 1980.

Sheppard, L.C. Computer control of the infusion of vasoactive drugs. *Ann. Biomed. Eng.*, 8:431–444, 1981.

Sheppard, L.C., N.T. Kouchoukos, J.F. Shotts, and F.D. Wallace. Regulation of mean arterial pressure by computer control of vasoactive agents in postoperative patients. In *Computers in Cardiology*, pages 91–94, Rotterdam, Netherlands, October 2–4 1975.

Sheppard, L.C. and B.McA. Sayers. Dynamic analysis of the blood pressure response to hypotensive agents studied in post surgical patients. *Comput. Biomed. Res.*, 10:237–246, 1977.

Sheppard, L.C., J.F. Shotts, N.T. Roberson, F.D. Wallace, and N.T. Kouchoukos. Computer controlled infusion of vasoactive drugs in post cardiac surgical patients. In *Proc. First Ann. Conf. IEEE Eng. in Medicine and Biology Soc.*, pages 280–284, Denver, Colorado, October 6-7 1979.

Shook, D.S., C. Mohtadi, and S.L. Shah. Identification for long range predictive control. *IEE Proc.-D*, 138(1):75–84, 1991.

Sirisena, H.R. and F.C. Teng. Multivariable pole-zero placement controller. *Int. J. Sys. Sci.*, 17(20):345–352, 1986.

Slate, J.B. *Model-based design of a controller for infusing sodium nitroprusside during postsurgical hypertension.* Ph.D. thesis, University of Wisconsin, 1980.

Slate, J.B. and L.C. Sheppard. Automatic control of blood pressure by drug infusion. *IEE Proc.-A*, 129(9), December 1982a.

Slate, J.B. and L.C. Sheppard. A model-based adaptive blood pressure controller.

In *IFAC Identification and System Parameter Estimation*, pages 1437–1442, 1982b.

Slate, J.B., L.C. Sheppard, V.C. Rideout, and E.H. Blackstone. A model for design of a blood pressure controller for hypertensive patients. In *5th IFAC Symposium on Identification and System Parameter Estimation*, pages 285–289, Darmstadt, Germany, September 1979.

Slate, J.B., L.C. Sheppard, V.C. Rideout, and E.H. Blackstone. Closed-loop nitroprusside infusion: modeling and control theory for clinical application. In *IEEE 1980 International Symposium on Circuits and Systems*, pages 482–488, 1980.

Sobel, K., H. Kaufman, and L. Mabius. Implicit adaptive control for a class of mimo systems. *IEEE Transactions on Aerospace and Electron System*, AES-18:576, 1982.

Steinmetz, M. Model reference adaptive control of blood pressure. M.Sc. thesis, Florida Atlantic University, 1987.

Stern, K.S., H.J. Chizeck, B.K. Walker, P.S. Krishnaprasad, P.J. Dauchot, and P.G. Katona. The self-tuning controller: comparison with human performance in the control of arterial pressure. *Annals of Biomedical Engineering*, 13:341–357, 1985.

Stern, K.S., B.K. Walker, and P.G. Katona. Automated blood pressure control using a self-tuning regulator. In *IEEE 1981 Frontiers of Engineering in Health Care*, pages 255–258, 1981.

Voss, G.I., H.J. Chizeck, and P.G. Katona. Self-tuning controller for drug delivery systems. *Int. J. Cont.*, 47(5):1507–1520, May 1988.

Voss, G.I., P.G. Katona, and H.J. Chizeck. Automated control of arterial pressure and cardiac output with nitroprusside and dobutamine in anesthetized dogs. In *Proc. American Control Conference*, pages 874–877, NY, 1986.

Voss, G.I., P.G. Katona, and H.J. Chizeck. Adaptive multivariable drug delivery: control of arterial pressure and cardiac output in anesthetized dogs. *IEEE Trans. Biomed. Eng.*, BME-34(8):617–625, August 1987.

Vozeh, S. and J.L. Steimer. Feedback control methods for drug dosage optimisation. *Clinical Pharmacokinetics*, 10:457–476, 1985.

Walker, B.K., T.L. Chia, K.S. Stern, and P.G. Katona. Parameter identification and adaptive control for blood pressure. In *IFAC Identification and System Parameter Estimation*, pages 1413–1418, Wash. D.C., 1982.

Wesseling, K.H., J.J. Settels, H.G. Walstra, H.J. VanEsch, and J.J.H. Donders. Baromodulation as the cause of short term blood pressure variability? In G. Alberi and Z. Bajzer, editors, *Proc. Intern. Conf. on Appl. Physics to Med Biol.*, Singapore, 1982. World Scientific Publishing.

Westenskow, D.R. Fundamentals of feedback control applied to microcomputer instrumentation design. *Int. J. Clin. Monit. Comput.*, 3:239–244, 1986.

Westenskow, D.R., L. Meline, and N.L. Pace. Controlled hypotension with sodium nitroprusside: anesthesiologist vesus computer. *J. Clin. Monit.*, 3:80–86, 1987.

Yamashita, Y., M. Suzuki, and K. Kambe. Fuzzy control of blood pressure by drug infusion. *Journal of Chemical Engineering of Japan*, 21(5):541–543, October 1988.

Ying, H., L. Sheppard, and D. Tucker. Expert-system-based fuzzy control of arterial pressure by drug infusion. *Medical Progress Through Technology*, 13:203–215, 1988.

Yu, C., R.J. Roy, H. Kaufman, and B.W. Bequette. Multiple-model adaptive predictive control of mean arterial pressure and cardiac output. Submitted to IEEE Trans on Biomedical Engineering, 1991.

Zhang, S., B. Hou, L. Peng, and G. Wang. Adaptive control for blood pressure with dahlin algorithm. In *Proc. of the 1988 IEEE International Conference on Systems, Man, and Cybernetics*, volume 1, pages 202–205. International Academic Publishers, 1988.

Chapter 6

Development of LRPC System for Mean Arterial Pressure Regulation[1]

6.1 Introduction

Methods for the automated control of mean arterial pressure (MAP) in postoperative patients have been investigated by many researchers. Some control systems based on feedback control theory including constraints werealso put into trials or routine clinical practice (Sheppard *et al.*, 1979; Rosenfeldt *et al.*, 1986; McNally *et al.*, 1977). However, improvements in the robustness of the closed-loop control algorithm and the system's adaptivity to a wide variety of patients of differing sensitivities are required. These changes are necessary because patients' response to commonly used vasodilators, e.g. sodium nitroprusside (SNP), are different and the response may even change within the same patient during the course of continuous drug infusion. Therefore systems having the ability of adapting to changes on-line are more desirable for arterial pressure control.

Many closed-loop adaptive control systems have been developed and tested.

---

[1] A version of this chapter has been published in the *International Journal of Adaptive Control and Signal Processing*: Constrained long-range adaptive predictive control of arterial blood pressure, Kwok, K.Y., R.K. Mutha, S.L. Shah, A.S. Clanachan, and B. Finegan, 5(6):363-374, 1991.

A number of them are of the self-tuning type. A (non-adaptive) minimum variance controller was implemented by Koivo *et al.* (1981a,b). Meline *et al.* (1985) describe the use of a self-tuning controller with a minimum variance control law. In order to combat the varying and unknown time delay problem, a pole-placement-based self-tuning method was attempted in a multiple-drug infusion system by Serna *et al.* (1983). Similar systems in the form of the Generalized Minimum Variance controller were tested by Arnsparger *et al.* (1983), Stern *et al.* (1985), Peng *et al.* (1988), and Millard *et al.* (1987,1988). A one-step ahead predictive controller for multiple-drug infusion was presented by McInnis and Deng (1985). The performance of an adaptive Dahlin algorithm approach was reported by Zhang *et al.* (1988). Recently, Voss *et al.* (1988) resolved the non-minimum phase problem in the multi-input/multi-output control of cardiac output and blood pressure control by using the Control Advance Moving Average Controller, a member of the class of extended horizon control technique proposed by Ydstie *et al.* (1985). Multiple-model adaptive control employing a weighted-average of the control actions from a bank of models was considered by He *et al.* (1986) and Martin *et al.* (1987).

The methods examined so far are one- or k-step ahead single point predictive controllers. This chapter describes a closed-loop adaptive control system based on the long range prediction strategy known as Generalized Predictive Control (GPC). The system was previously tested in the control of MAP on anesthetized dogs (Kwok *et al.*, 1990). The philosophy of this GPC approach is that rather than computing a control action based solely on the prediction of a single point in the future, a series of control actions in the future are considered by having a controlled output follow an entire trajectory of desired blood pressure. As a result, the whole control scheme is more robust than many other single-point prediction control algorithms reported in the literature. The algorithm with finite horizons was shown to produce stabilizing controllers (Clarke and Mohtadi, 1989). Also, because of the long-range output prediction, processes with non-minimum phase characteristics and varying time delay can also be easily accommodated. The detailed theoretical framework plus demonstrations of the above mentioned properties are given in Clarke *et al.* (1987a,b), Mohtadi (1987, 1988), McIntosh (1988), and McIntosh *et al.* (1988).

Recently, Yu *et al.* (1991) also included the GPC algorithm into their studies in the control of MAP and cardiac output. Their control strategy used a multiple model adaptive controller with a GPC controller for each model. However, the present work describes a true model-adaptive controller rather than a controller with a limited number of fixed models to be switched between. This work also includes results of both constrained and unconstrained control obtained from experiments on anesthetized dogs.

6.2 Controller Design

6.2.1 Model-Based Predictive Control

Slate *et al.* (1980) developed a model relating MAP to SNP infusion based on patient data together with the knowledge of physiological concepts plus pharmacological action of the drug. The model is a first-order transfer function with two time delay terms,

$$\frac{\Delta P(s)}{I(s)} = \frac{Ke^{-\tau_1 s}(1 + \alpha e^{-\tau_2 s})}{(1 + \tau s)} \tag{6.1}$$

where ΔP corresponds to the drop of MAP induced by SNP infusion rate I, K is the sensitivity of the patient to the drug, τ_1 and τ_2 represent the time delays due to transport lag and recirculation time, α is the recirculation fraction, and τ is the time constant.

Their studies also showed a large range of variations in the model parameters from patient to patient or even for the same patient during the course of continuous infusion. The studies contained in this chapter assume that the process which represents MAP response to SNP infusion is described by an autoregressive, integrated moving average model with an auxiliary input (ARIMAX),

$$\hat{A}(q^{-1})y(t) = \hat{B}(q^{-1})u(t-1) + T(q^{-1})\xi(t)\frac{1}{\Delta} \tag{6.2}$$

where $\hat{A}, \hat{B}$, and T are polynomials in the backward shift operator q^{-1} which is dropped in the sequel for brevity. $y(t)$ is the MAP. $u(t-1)$ represents the drug infusion rate with a zeroth-order hold. $\xi(t)$ is an uncorrelated random sequence with zero mean and Δ is the difference operator $1 - q^{-1}$.

By making use of the Diophantine identity,

$$T = E_j\hat{A}\Delta + q^{-j}F_j \tag{6.3}$$

which uniquely defines E_j and F_j, a future process output response model is obtained as follows,

$$\hat{y}(t+j \mid t) = G_j\Delta u(t+j-1) + f(t+j \mid t) \tag{6.4}$$

where $G_j = E_j\hat{B}$ with its coefficients corresponding to step-response model coefficients and $f(t+j \mid t)$ is the future "unforced" process output prediction term. Both polynomials $\hat{A}$ and $\hat{B}$ are obtained from an on-line identifier. The observer polynomial T is usually a user-defined noise model.

In the GPC strategy, the control objective is to keep a trajectory of predicted future process output as close to the setpoint as possible by generating a future control trajectory. The objective is defined by the cost function,

$$J(n1, n2, nu) = \sum_{j=n1}^{n2}[y(t+j \mid t) - w(t+j)]^2 + \sum_{j=1}^{nu}\lambda(j)[\Delta u(t+j-1)]^2 \tag{6.5}$$

where

- $n1$ is the minimum output prediction horizon,
- $n2$ is the maximum output prediction horizon,
- nu is the control action horizon, such that $\Delta u(t+i) = 0, i \geq \mathrm{nu}$,
- $\lambda(j)$ is a weighting sequence.

The following control law minimizes eqn. 6.5:

$$\mathbf{U} = (\mathbf{G^T G} + \mathbf{\Lambda})^{-1}\mathbf{G^T}(\mathbf{W} - \mathbf{F}) \tag{6.6}$$

where

$$\mathbf{G} = \begin{bmatrix} g_{n1-1} & \cdots & g_0 & 0 & 0 & \cdots & 0 \\ g_{n1} & \cdots & g_1 & g_0 & 0 & \cdots & 0 \\ \vdots & \ddots & & & \ddots & & \vdots \\ \vdots & & \ddots & & & & g_0 \\ \vdots & \cdots & & & & & \vdots \\ g_{n2-1} & g_{n2-2} & \cdots & & & & g_{n2-nu} \end{bmatrix} \tag{6.7}$$

$$\Lambda = \begin{bmatrix} \lambda(1) & 0 & \cdots & 0 \\ 0 & \lambda(2) & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda(nu) \end{bmatrix} \tag{6.8}$$

$$\mathbf{W} = \left[\ w(t+n1 \mid t) \quad w(t+n1+1 \mid t) \quad \cdots \quad w(t+n2 \mid t)\ \right]^T \tag{6.9}$$

$$\mathbf{F} = \left[\ f(t+n1 \mid t) \quad f(t+n1+1 \mid t) \quad \cdots \quad f(t+n2 \mid t)\ \right]^T \tag{6.10}$$

Since the algorithm is functioning in a receding horizon manner, only the first element of eqn. 6.6 is implemented and a new vector **U** is calculated at the next sample interval. Details of the GPC derivation, its variations and properties are documented elsewhere (Clarke *et al.*, 1987a,b; Clarke and Mohtadi, 1989; Mohtadi, 1988; McIntosh *et al.*, 1988; McIntosh, 1988).

6.2.2 Constrained Optimization

The control action implemented on a physical system has limitations due to saturation of control elements; for example, any valve can only open between the range of 0 % and 100 % or the flow rate of a heating fluid cannot go negative. In the control of MAP, additional physical and physiological constraints as follows also need to be observed,

- Drug infusion limit: a maximum infusion of SNP is to be fixed to prevent high SNP concentration in the body.
- Incremental change of arterial blood pressure: the rate of change of arterial blood pressure should be limited to avoid complications caused by sudden changes in blood pressure.
- Bounds on the range of arterial blood pressure: arterial blood pressure should not drop or rise beyond acceptable limits.

Because of the predictive ability of the GPC algorithm, an optimal solution that minimizes eqn. 6.5 subject to the above constraints can be found. Since all constraints on arterial blood pressure can always be mapped into constraints on drug infusion via the prediction model, only the following two constraints will be considered in the minimization of eqn. 6.5,

the rate constraint:

$$-\alpha_u \leq \Delta u \leq \alpha_u \tag{6.11}$$

and the amplitude constraint:

$$\rho_1 \leq u \leq \rho_2 \tag{6.12}$$

Constraint 6.11 corresponds to the maximum allowable change of infusion at each interval in either direction. Constraint 6.12 corresponds to the minimum and maximum dose rate for a particular patient. ρ_1 is usually zero, corresponding to the minimum possible infusion rate. It is considered important in constrained optimization because no negative infusion rate can be implemented in practice. ρ_2 is a variable dependent on the subject's weight and the total duration of SNP administration. For the application of SNP on patients, the maximal dose should not exceed 800 μg min^{-1} (Goodman and Gilman, 1975).

6.2.3 Process Identification

The coefficients in polynomials $\hat{A}$ and $\hat{B}$ are estimated on-line using a recursive least square algorithm with a variable forgetting factor. Since this algorithm has been used extensively, the reader is referred to Seborg *et al.* (1986) for details.

6.2.4 Implementation

A second order ARIMAX model was assumed to represent the response of MAP to SNP infusion. The varying time-delay was accommodated by using an extended $\hat{B}(q^{-1})$ polynomial. A total of seven parameters were identified by the recursive least squares algorithm incorporated with UD factorization and variable forgetting. The sample time was selected to be 10 seconds. High-frequency noise was filtered by the observer polynomial,

$$T = 1 - 0.8q^{-1} \tag{6.13}$$

before the data were sampled by the identification routine.

All control actions were penalized by a control weight of $\lambda = 0.001$ mainly to ensure that the matrix inversion in eqn. 6.6 was non-singular. The initial values for minimum output horizon ($n1$), maximum output horizon ($n2$) and control horizon (nu) were selected to be 2, 10, and 1 respectively, as these were suggested default values at commissioning stage.

An analytical solution to the constrained GPC problem was developed for some simple SISO ($nu \leq 2$) (Tsang and Clarke, 1988) and MIMO ($nu = 1$) cases (Mutha, 1991). Experimental evaluation of such analytically solved constrained GPC problems was documented by Mutha (1991). For the purpose of devising and evaluating a constrained GPC controller for a general class of rate and amplitude constraints on the inputs and outputs, the constrained algorithm is implemented using the optimization package QPSOL from Stanford University which searches for an optimal solution in the quadratic space iteratively. Mixed-language programming is necessary because the QPSOL package is in FORTRAN whereas all in-house routines are written in C. The performance of the controller in the presence of constraints was tested on a pilot-scale continuous stirred-tank heater and found to have satisfactory disturbance rejection as compared to the unconstrained GPC controller (Mutha, 1991).

To avoid the possibility of over-dosing the subject in the event of a sudden drop in MAP, the program has a safety-check routine which validates the control signal before sending the signal to the infusion pump. Should the MAP fall below a certain critical user-specified limit, the routine would disregard the negative incremental constraint on Δu and reduce SNP infusion in an aggressive manner.

6.3 System Description

6.3.1 Equipment

A schematic diagram of the experimental setup is shown in fig. 6.1. An HP 78200 series monitoring system was used to measure the physiological parameters. The arterial pressure wave was measured via a carotid using a pressure transducer. The signal was amplified and the systolic as well as the diastolic pressure detected. The MAP was calculated and made available both on a digital display board and as a 0

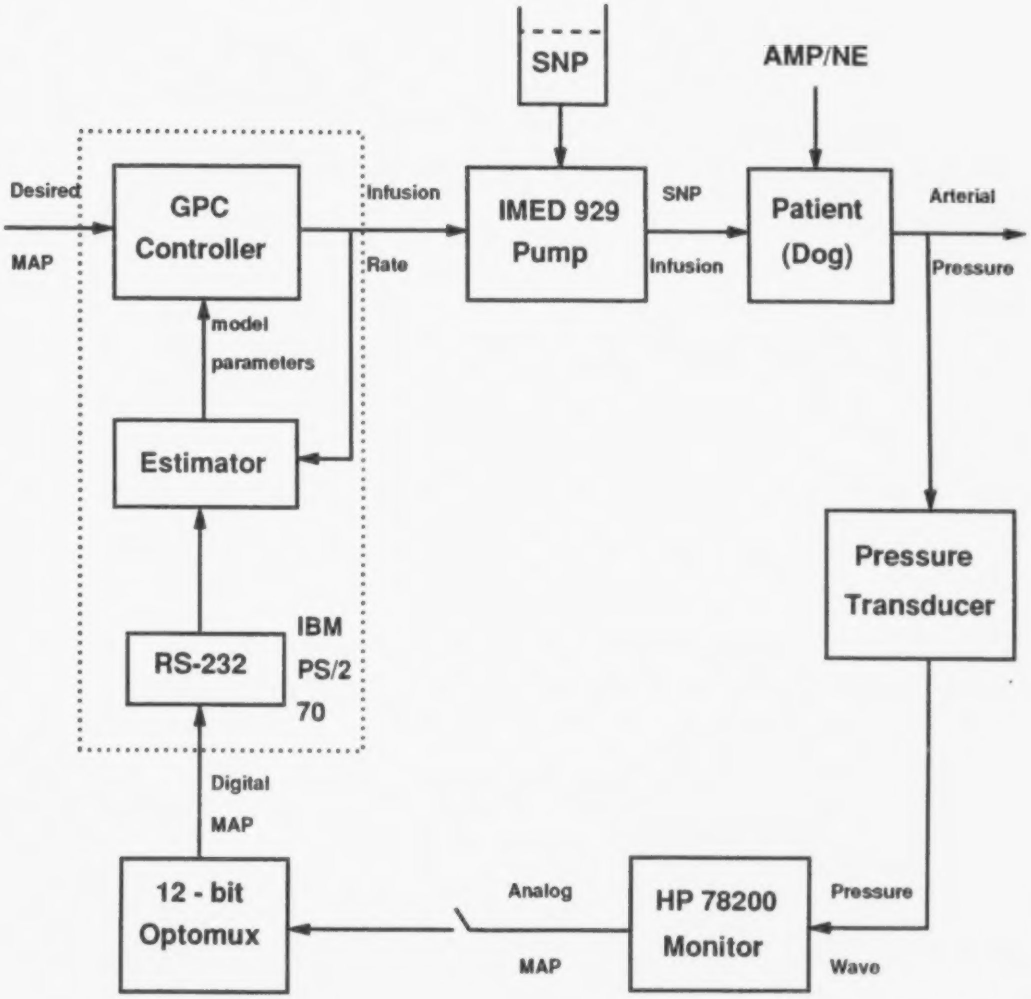
SNP
AMP/NE
Desired
MAP
GPC
Controller
Infusion
Rate
IMED 929
Pump
SNP
Infusion
Patient
(Dog)
Arterial
Pressure
model
parameters
Estimator
Pressure
Transducer
RS-232
IBM
PS/2
70
Digital
MAP
12 - bit
Optomux
Analog
MAP
HP 78200
Monitor
Pressure
Wave

Figure 6.1: Block diagram of the adaptive blood pressure control system

to 3 V analog signal (corresponding to 0 to 300 mmHg). The blood pressure signal was sampled by an Optomux system every 10 seconds, and converted from an analog signal to a 12-bit digital signal. An IBM Model-70 PS/2 personal computer running under the QNX operating system was employed for the control calculations and data acquisition. It receives the 12-bit blood pressure digital signal from the Optomux system via an RS-232 interface. After an appropriate control action is calculated, the signal is sent to an IMED[2] 929 computer-enabled drug infusion pump which has a capacity from 0 to 1599 ml hr^{-1}. The infusion rate is updated at every sample time. Two drugs were used in the closed-loop control experiment, namely SNP, a vasodilator, and norepinephrine, a vasoconstrictor. The former was used as the manipulated variable to decrease the blood pressure and the latter as a disturbance to increase blood pressure. Both were delivered via a triple-lumen catheter positioned in the femoral vein.

6.3.2 Software

In an ideal automatic control environment, continuous monitoring and control in real time should be maintained with minimal interruption from other operations such as changing setpoint or control parameters. The continuity is also essential in adaptive blood pressure control because blood pressure as well as the adaptive property of the controller cannot afford frequent interruption or intermittent termination. Therefore, a real-time, multi-tasking environment, in which the computer can respond to the operator's requests while maintaining control was considered essential in our development of the blood pressure controller. In order to achieve this environment with a flexible and fast program-operator communication interface, all software programs were written in the **C** language under the QNX operating system, which is a UNIXbased real-time, multi-tasking operating system. A MULTI-purpose CONtrol system package (MULTICON) has been developed at the Department of Chemical Engineering at the University of Alberta. It provides a flexible environment for control system management under which the timing, sampling via the Optomux system, scheduling of tasks, and real-time graphic display of data are performed in a multi-tasking and real-time manner. A schematic of the software system is shown in fig. 6.2. Three user tasks are supervised by MULTICON: a recursive least squares identification routine, the GPC control routine, and the pump driver routine. Because of a relatively large computational load and large memory requirements, the GPC routine was executed with the dynamic allocation of computer memory space. Since MULTICON is multi-tasking, the operator can then change the setpoint and control parameters on-line without disturbing closed-loop control.

---

[2] IMED Corporation, San Diego, California

6.4 Experimental Studies

6.4.1 Methods

Following ethics approval, experiments were performed on anesthetized dogs (20-25 kg) in the laboratory of the Department of Anaesthesia/Pharmacology at the University of Alberta. Anesthesia was induced by 30 mg kg^{-1} sodium pentobarbital and maintained by a continuous infusion of pentobarbital (3 mg $kg^{-1}hr^{-1}$), fentanyl (20 μg $kg^{-1}hr^{-1}$) and pancuronium (100 μg $kg^{-1}hr^{-1}$). SNP at a concentration of 200 μg ml^{-1} was used as the manipulated variable.

The commissioning period of blood pressure control consists of two steps: an open-loop pseudo-random binary sequence (PRBS) with an amplitude decided by medical personnel and a closed-loop setpoint change. The PRBS serves as an initial probing signal for the control system and initialization of model parameters. The

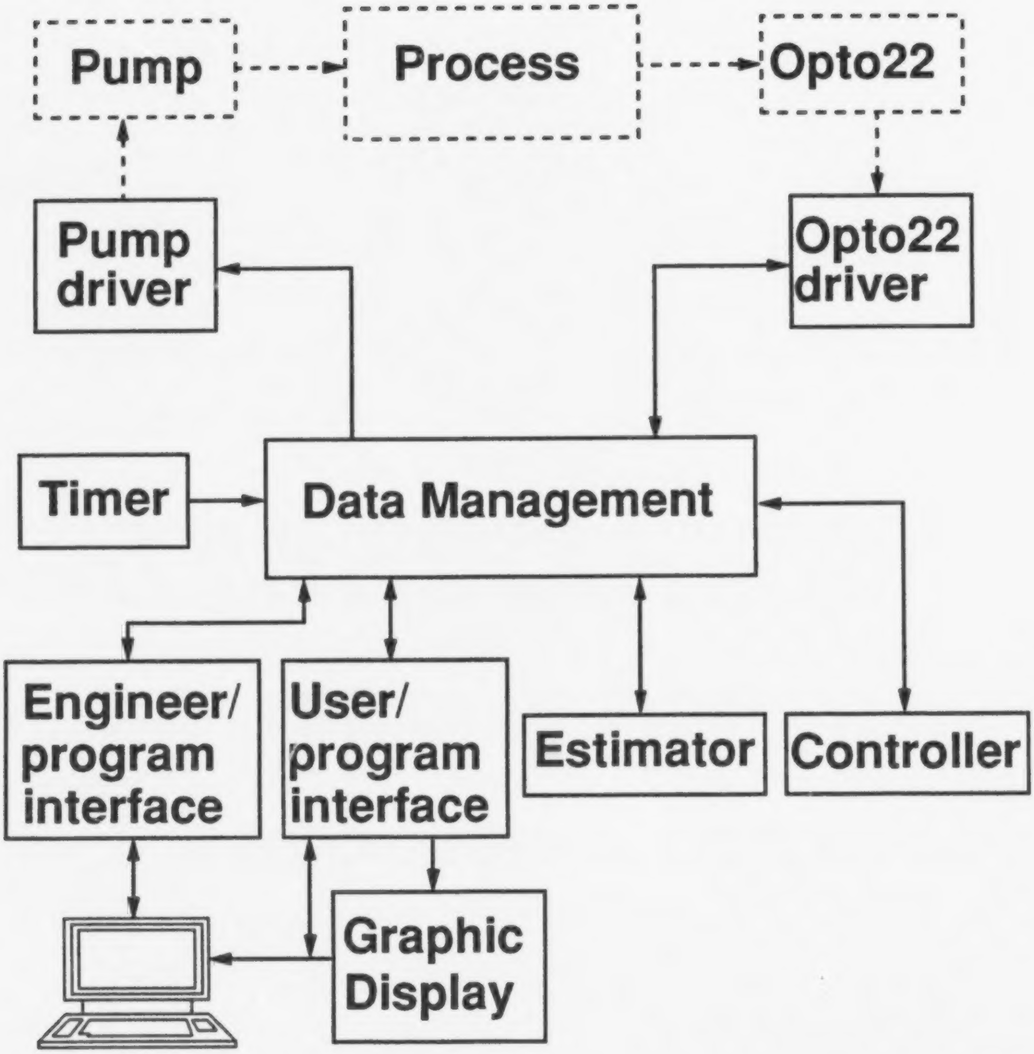
Pump
Process
Opto22
Pump driver
Opto22 driver
Timer
Data Management
Engineer/ program interface
User/ program interface
Estimator
Controller
Graphic Display

Figure 6.2: Block digram of the MULTICON real-time, multi-tasking control program

subsequent setpoint change was to lower the blood pressure from a hypertensive state to a lower level. The performance of the controller was then evaluated upon a series of unpredictable disturbances introduced artificially by infusion of other vasoactive drugs. Norepinephrine (NE) which causes vasoconstriction was used to induce hypertension. Adenosine mono-phosphate (AMP) which has a similar effect as SNP was applied to lower blood pressure. Both drugs were separately introduced in continuous infusion as well as in bolus.

6.4.2 Results

The procedure for evaluating the controller on dogs included administering disturbance drugs during an open-loop (OL) control run, a closed-loop constrained control run and a closed-loop unconstrained control run. The purpose of the open-loop controlled run was to illustrate the severity of the effect caused by disturbance drugs on the subject while the disturbance was not being compensated by any corrective actions. During the period of this OL control run, SNP was continuously infused at a fixed rate to the subject so as to generate a hypotensive state close to the one during a closed-loop control run. The two closed-loop control runs were separated by the OL control run. Replicate experiments were done on five dogs which were anesthetized during the experiment. A typical set of results is included in this chapter to illustrate the performance of both constrained and unconstrained GPC.

Before all these runs were performed on the same subject, open-loop step responses of the MAP in response to SNP infusions were first recorded and a typical sample is plotted in fig. 6.3. SNP infusions were increased from 2.67 μg kg^{-1}min^{-1} to 20 μg kg^{-1}min^{-1} in 6-minute intervals. Fig. 6.4 shows the unconstrained control of MAP. A PRBS with a period of 15 samples started the initialization of parameters. Following a small drop in blood pressure, a setpoint change was made from the current

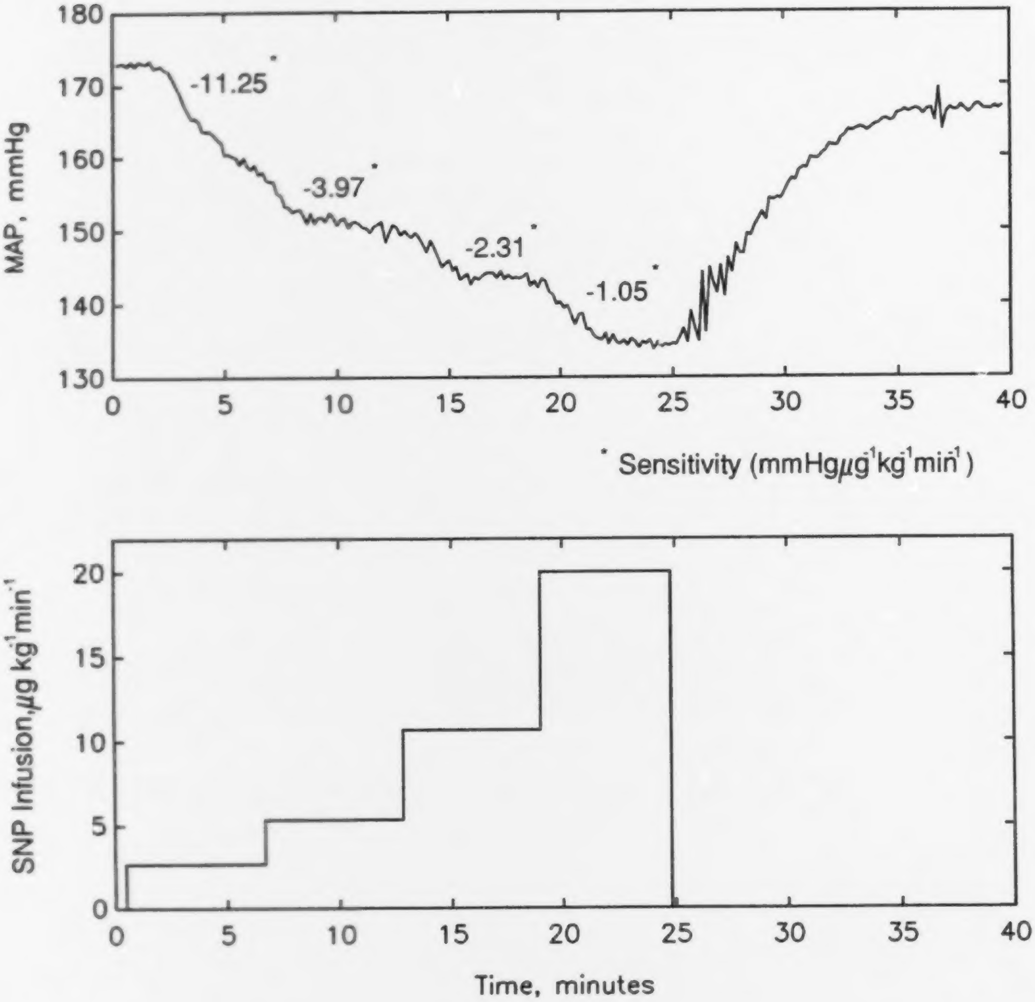

Figure 6.3: Open-loop step responses of MAP

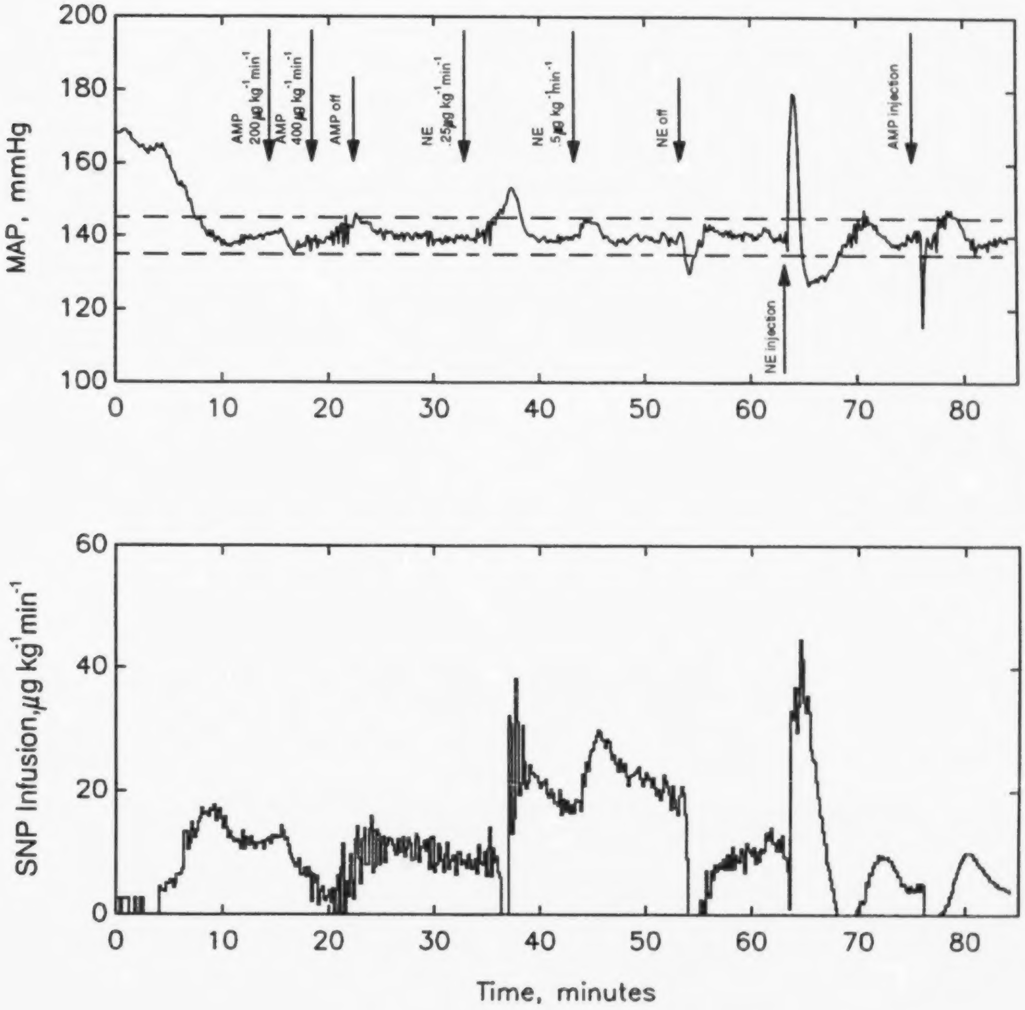

Figure 6.4: Trajectory of MAP in response to disturbances under unconstrained GPC control

Time(min.)	Event
0	Setpoint change from 170 to 140 mmHg
15	Adenosine monophosphate at 200 μg kg^{-1}min^{-1}
18	Adenosine monophosphate at 400 μg kg^{-1}min^{-1}
21	Adenosine monophosphate infusion off
33.3	Norepinephrine at 0.25 μg kg^{-1}min^{-1}
43.3	Norepinephrine at 0.5 μg kg^{-1}min^{-1}
53.3	Norepinephrine infusion off
63.3	Norepinephrine injection, 1 μg kg^{-1}
76	Adenosine monophosphate injection, 800 μg kg^{-1}

Table 6.1: Challenge procedure during closed-loop control

MAP to 140 mmHg. Continuous infusion of AMP at 200 μg kg^{-1}min^{-1} started at t=15 min and was doubled after 3 minutes. After AMP stopped and the MAP returned to the setpoint, continuous infusion of NE started at t=33.3 min at a rate of 0.25 μg kg^{-1}min^{-1}. Because the effect of NE was more drastic than that of AMP, NE infusion lasted 10 minutes before it was doubled to 0.5 μg kg^{-1}min^{-1}. The subsequent disturbances were NE and AMP injections in bolus consecutively. Table 6.1 lays out the details of the challenge procedure in sequence.

The OL control run is showed in fig. 6.5. It should be noted that the MAP was lowered to 140 mmHg, the same hypotensive level as in the previous run. Manual adjustment of SNP infusion was required from time to time in order to keep the MAP at around the setpoint, while no disturbance drug was introduced.

The constrained control run in fig. 6.6 allows a comparison with the unconstrained one in fig. 6.4. The schedule of disturbances for both runs were similar in sequence and the same in drug concentrations (see table 6.1). However, the level of sensitivity of the subject to SNP changed after the subject had been infused with SNP on and off for several hours. The starting baseline pressure in fig. 6.6 was already lower than that in fig. 6.4.

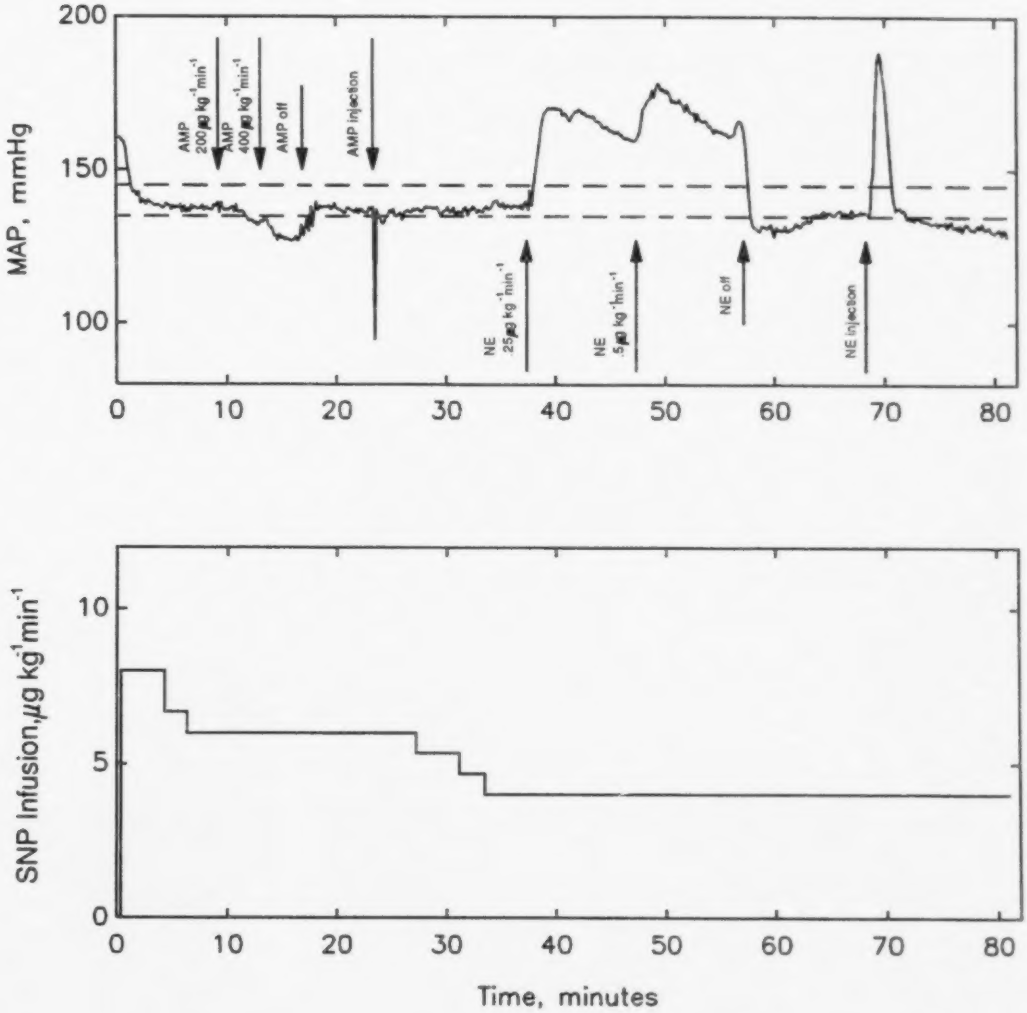

Figure 6.5: Trajectory of MAP in response to disturbances at open-loop fixed SNP infusion rate

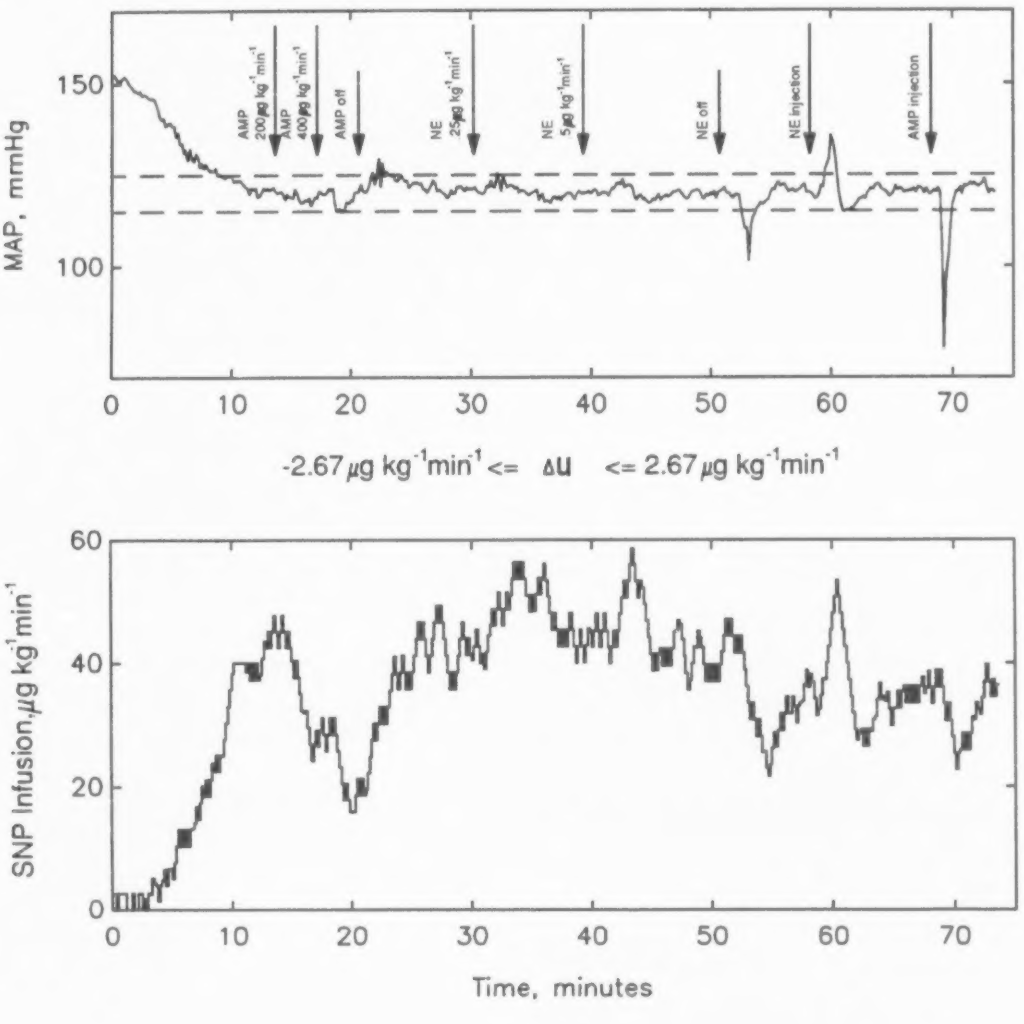

Figure 6.6: Trajectory of MAP in response to disturbances under constrained GPC control

6.5 Discussion

Non-linear dynamics in chemical processes are a great challenge to satisfactory control. This challenge is no exception in biological processes. The step responses in fig. 6.3 clearly present the degree of non-linearity in controlling blood pressure. The sensitivity of the subject (*i.e.* the negative gain of the process) changes from -11.25 mmHg $\mu g^{-1} kg^{-1} min^{-1}$ at the initial infusion of 2.67 μg $kg^{-1} min^{-1}$ to -1.05 mmHg $\mu g^{-1} kg^{-1} min^{-1}$. The amount of SNP required was 10 times more to cause only a 10 mmHg drop at a lower blood pressure level than that at a high blood pressure level. This increase in SNP infusion was obviously due to the fact that the subject had built up a certain tolerance to the effect of SNP and other reflexes of the body started to counteract the sudden drop in blood pressure. The asymmetric behavior of blood pressure was observed when the SNP infusion ceased and the MAP did not return to the baseline. It demonstrated in fig. 6.3 that process dynamics was changing during closed-loop control; therefore an adaptive mechanism would play an important role in achieving the control objective.

The artificial disturbances introduced by other drug interventions were designed to cause heavy hypertension and mild hypotension. If excessive hypotension was induced by AMP to a level which no longer requires SNP therapy, the controller would simply set itself to 0 infusion rate during AMP infusion, and its response to hypotension could not be observed completely. Therefore, a strong dosage of NE was used to cause about 40 mmHg rise and a mild dosage of AMP to cause about 20 mmHg drop (see fig. 6.5).

Consider the unconstrained GPC control in fig. 6.4. The introduction of continuous AMP infusion caused a gradual decrease in SNP infusion. When comparing the AMP effect between closed-loop control and the open-loop run (fig. 6.5), one will discover the similar shape between the SNP drop and MAP drop in figures 6.4 and 6.5 respectively. The effect of NE on blood pressure was counteracted by the sharp increase in SNP infusion. However, a non-minimum phase behavior from the controller was observed at t=35.8 min, caused by a right-hand-plane zero in the process model.

The NE injection at t=63.3 min in fig. 6.4 induced a jump of 39 mmHg in MAP. In response to this initial jump, the controller took an aggressive action in bringing down the MAP. Since the NE effect was quick in wearing off, the subject was overdosed by SNP causing a 15 mmHg overshoot. Another overshoot was observed after the AMP injection at t=76 min. The overdose did not appear to be a serious problem in the case of constrained GPC control (see fig. 6.6). The reason is obviously due to the constraints which limit the aggressive infusion of SNP.

The results shown in fig. 6.6 were obtained during the last stage of the experiment when the subject had built up a stronger tolerance to SNP. The baseline in this figure is lower than that in fig. 6.4. Because more SNP was required to decrease MAP at a lower hypotensive state (which is in agreement with the result in fig. 6.3), the controller encountered the amplitude constraint (at t=10.3 min) which was subsequently relaxed to 60 μg kg^{-1}min^{-1}. The same dosage of NE used in the constrained GPC case appears to be less effective when applied to the unconstrained case. This observation is indicated by the momentarily mild increase in SNP during the period between t=31 min and t=52 min. On the other hand, AMP caused a larger change in MAP. Nevertheless, the controller managed to maintain satisfactory control during the whole run.

6.6 Conclusions

A multi-step adaptive predictive controller has been experimentally evaluated for control of blood pressure. The purpose of this controller is to control the mean arterial pressure under hypertensive conditions using sodium nitroprusside to maintain the pressure at a desired lower level. The development of the controller makes use of a real-time, multi-tasking software environment which allows operators to make changes on-line.

Norepinephrine and adenosine mono-phosphate were employed as disturbance drugs to induce hypertension and hypotension respectively. The disturbances were designed to simulate extreme scenarios which a post-operative patient might encounter. During the whole course of the experiment on the same subject, both unconstrained and constrained Generalized Predictive Control performed satisfactorily in keeping the mean arterial pressure close to the setpoint. This controller design using Generalized Predictive Control is believed to surpass single-point predictive control algorithms by providing more robustness in an environment in which unpredictable disturbances abound.

Bibliography

Arnsparger, J.M., B.C. McInnis, J.R. Glover, and N.A. Normann. Adaptive control of blood pressure. *IEEE Trans. Biomed. Eng.*, BME-30(3):168–176, March 1983.

Clarke, D.W. and C. Mohtadi. Properties of Generalized Predictive Control. *Automatica*, 25(6):859–875, 1989.

Clarke, D.W., C. Mohtadi, and P.S. Tuffs. Generalised predictive control – Part I. The basic algorithm. *Automatica*, 23(2):137–148, 1987a.

Clarke, D.W., C. Mohtadi, and P.S. Tuffs. Generalised predictive control – Part II. Extensions and interpretations. *Automatica*, 23(2):149–160, 1987b.

Goodman, L.S. and A. Gilman. *The Pharmacological Basis of Therapeutics*. Macmillan, New York, 5 edition, 1975.

He, W.G., H. Kaufman, and R. Roy. Multiple model adaptive control procedure for blood pressure control. *IEEE Trans. Biomed. Eng.*, BME-33(1):10–19, January 1986.

Koivo, A.J. Microprocessor-based controller for pharmacodynamic applications. *IEEE Trans. Auto. Cont.*, AC-26:1208–1212, 1981.

Koivo, A.J., D. Larnard, and R. Gray. Digital control of mean arterial blood pressure in dogs by injecting a vasodilator drug. *Annals of Biomedical Engineering*, 9:185–197, 1981.

Kwok, K.Y., A.S. Clanachan, B. Finegan, and S.L. Shah. Long range predictive control of arterial blood pressure. In *Proc. 29th IEEE Conference on Decision and Control*, pages 2800–2805, Honolulu, Hawaii, December 5-7 1990.

Martin, J.F., A.M. Schneider, and N.TY Smith. Multiple-model adaptive control of blood pressure using sodium nitroprusside. *IEEE Trans. Biomed. Eng.*, BME-34(8):603–611, August 1987.

McInnis, B.C. and L.Z. Deng. Automatic control of blood pressures with multiple drug inputs. *Annals of Biomedical Engineering*, 13:217–225, 1985.

McIntosh, A.R. Performance and tuning of adaptive generalized predictive control. M.Sc. thesis, University of Alberta, 1988.

McIntosh, A.R., S.L. Shah, and D.G. Fisher. Experimental evaluation of adaptive control in the presence of disturbances and model-plant mismatch. In S.L. Shah and G. Dumont, editors, *Adaptive Control Strategies for Industrial Use*, volume 137 of *Springer-Verlag Lecture Notes in Control and Information Sciences*, pages 145–174. Springer-Verlag, New York, 1988.

McNally, R.T., K. Engelman, A. Noordergraaf, and M. Edwards. A device for the precise regulation of blood pressure in patients during surgery and hypertensive crisis. In *Proc. San Diego Biomedical Symposium*, volume 16, pages 419–424, New York, 1977. Academic Press.

Meline, L.J., D.R. Westenskow, N.L. Pace, and M.N. Bodily. Computer-controlled regulation of sodium nitroprusside infusion. *Anesth. Analg.*, 64:38–42, 1985.

Millard, R.K., P. Hutton, E. Pereira, and C. Prys-roberts. On using a self-tuning controller for blood pressure regulation during surgery in man. *Comput. Biol. Med.*, 17(1):1–18, 1987.

Millard, R.K., C.R. Monk, and C. Prys-Roberts. Surgery using a volatile anaesthetic. *IEE Proc.-D*, 135(2):95–105, March 1988.

Mohtadi, C. *Advanced Self-Tuning Algorithms.* D.Phil thesis, University of Oxford, 1987.

Mohtadi, C. On the role of prefiltering in parameter estimation and control. In S.L. Shah and G. Dumont, editors, *Adaptive Control Strategies for Industrial Use*, volume 137 of *Springer-Verlag Lecture Notes in Control and Information Sciences*, pages 121–144. Springer-Verlag, New York, 1988.

Mutha, R.K. Constrained long range predictive control. M.Sc. thesis, University of Alberta, 1990.

Peng, L., G. Wang, S. Zhang, and H. Hou. An adaptive controller for blood pressure. In *8th IFAC/IFORS symp. on Ident. syst. para. est.*, volume 3, pages 1835–1839, Beijing, China, 1988.

Rosenfeldt, F.L., V. Chang, M. Grigg, S. Parker, R. Cearns, M. Rabinov, and W.G. Xu. A closed loop microprocessor controller for treatment of hypertension after cardiac surgery. *Anaesth. Intens. Care*, 14:158–162, 1986.

Seborg, D.E., T.F. Edgar, and S.L. Shah. Adaptive control strategies for process control: a survey. *A.I.Ch.E. Journal*, 32:881–913, 1986.

Serna, V., R. Roy, and H. Kaufman. Adaptive control of multiple drug infusions. In *Proc. American Control Conference*, pages 22–26, San Francisco, CA, 1983.

Sheppard, L.C., J.F. Shotts, N.T. Roberson, F.D. Wallace, and N.T. Kouchoukos. Computer controlled infusion of vasoactive drugs in post cardiac surgical patients. In *Proc. First Ann. Conf. IEEE Eng. in Medicine and Biology Soc.*, pages 280–284, Denver, Colorado, October 6-7 1979.

Slate, J.B., L.C. Sheppard, V.C. Rideout, and E.H. Blackstone. Closed-loop nitroprusside infusion: modeling and control theory for clinical application. In *IEEE 1980 International Symposium on Circuits and Systems*, pages 482–488, 1980.

Stern, K.S., H.J. Chizeck, B.K. Walker, P.S. Krishnaprasad, P.J. Dauchot, and P.G. Katona. The self-tuning controller: comparison with human performance in the control of arterial pressure. *Annals of Biomedical Engineering*, 13:341–357, 1985.

Tsang, T.T.C. and D.W. Clarke. Generalized predictive control with input constraints. *IEE Proc.-D*, 135(6):451–460, 1988.

Voss, G.I., H.J. Chizeck, and P.G. Katona. Self-tuning controller for drug delivery systems. *Int. J. Cont.*, 47(5):1507–1520, May 1988.

Ydstie, B.E., L.S. Kershenbaum, and R.W.H. Sargent. Theory and application of an extended horizon self-tuning controller. *A.I.Ch.E. Journal*, 31(11):1771–1780, 1985.

Yu, C., R.J. Roy, H. Kaufman, and B.W. Bequette. Multiple-model adaptive predictive control of mean arterial pressure and cardiac output. Submitted to IEEE Trans on Biomedical Engineering, 1991.

Zhang, S., B. Hou, L. Peng, and G. Wang. Adaptive control for blood pressure with dahlin algorithm. In *Proc. of the 1988 IEEE International Conference on Systems, Man, and Cybernetics*, volume 1, pages 202–205. International Academic Publishers, 1988.

Chapter 7

Evaluation of LRPC System for Mean Arterial Pressure Regulation[1]

7.1 Introduction

A number of closed-loop control systems have been investigated and developed for the regulation of physiological variables by automatic administration of therapeutic agents, especially in the control of postoperative mean arterial pressure (MAP). It is concluded from several surveys of these developments that control strategies ranging from simple proportional-integral-derivative (PID) based feedback control to model-based adaptive predictive control have been applied (Katona, 1982; Linkens, 1984; Linkens and Hacisalihzade, 1990). The main driving force behind the use of more advanced control strategies was the recognition of the variability of patients' sensitivities to drug administration which necessitate the adaptation and self-tuning control ability of an automated physiological controller.

Sheppard *et al.* (1975) conducted most of the pioneering work in the control of MAP by the infusion of sodium nitroprusside (SNP), a commonly used vasodilator.

---

[1] A version of this chapter is to be presented as an invited paper at the IFAC International Symposium on Adaptive Control and Signal Processing: Evaluation of a long-range adaptive predictive controller for computerized drug delivery systems, Kwok, K.Y., S.L. Shah, A.S. Clanachan, and B.A. Finegan, 1992.

His subsequent work and modification of the original PID controller have led to the postulation of a well-known MAP model responding to SNP infusion (Slate *et al.*, 1980; Slate and Sheppard, 1982b) and routine clinical practice of such a controller on postoperative patients (Slate and Sheppard, 1982a). Other PID-based MAP controllers were evaluated in different scenarios (de Asla *et al.*, 1985; Meline *et al.*, 1986; Westenskow *et al.*, 1987; Reid and Kenny, 1987; Cosgrove III *et al.*, 1989; Colvin and Kenny, 1989).

Because of the varying nature of patients sensitivities to SNP, more advanced control strategies with predictive ability and a certain degree of adaptivity have been used. Those strategies with extensive evaluations include model reference adaptive control (Kaufman *et al.*, 1984), minimum variance and generalized minimum variance control (Stern *et al.*, 1985; Meline *et al.*, 1985; Millard *et al.*, 1987; Millard *et al.*, 1988), as well as control advance moving average control (Voss *et al.*, 1988). Their control objectives are all based on single-step ahead prediction, requiring proper tuning in order to avoid oscillations and instability. Recently, a class of long-range predictive controllers that provide excellent robustness have been investigated for possible application in the control of various physiological parameters (Linkens *et al.*, 1991). A study by Yu *et al.* (1991) utilized the generalized predictive control (GPC) algorithm (Clarke *et al.*, 1987) in their multiple model adaptive controller. However, the adaptation is limited to switching between a bank of pre-specified models which are applicable only for MAP control.

The present work evaluates a computerized drug delivery system which is truly model-adaptive. The control strategy combines GPC with a terminal matching condition so that the control objective involves a trajectory of future output predictions and the steady-state response (see Chapters 2 and 3). The estimation algorithm is based on a control-relevant long-range identification strategy which is compatible with the overall control objective (Shook *et al.*, 1991). This study compares the model-adaptive system's performance in the regulation of MAP with no computer control. A challenge protocol was designed to provide a fair comparison between computer-controlled and uncontrolled performance by obtaining the results at the same hypotensive state.

7.2 Process Control Strategy

7.2.1 Control Scheme

The control algorithm is a member of the class of long-range predictive control methods. It is based on the generalized predictive control (GPC) law (Clarke *et al.*, 1987) with the incorporation of a terminal matching condition, namely steady-state error weighting. Details of the control law derivation can be found in Chapter 2 and summarized in the following discussion:

The idea of applying GPC with steady-state error weighting to MAP control is to maintain the MAP at the desired target over a finite future horizon and also at steady state. Let a discrete physiological model ($\frac{\hat{A}(q^{-1})}{\hat{B}(q^{-1})}$) relating the effect of past SNP infusion ($u(\cdot)$) to MAP ($y(\cdot)$) be represented by the following time series model:

$$\hat{A}(q^{-1})y(t) = \hat{B}(q^{-1})u(t-k) + T(q^{-1})\frac{\xi(t)}{1-q^{-1}} \tag{7.1}$$

where $T(q^{-1})\frac{\xi(\cdot)}{1-q^{-1}}$ is the noise model and k is the time delay. Then a trajectory of future MAP changes can be related to a trajectory of future SNP infusions by shifting the time index in eqn. 7.1. The optimal future SNP infusion is found by minimizing the squares of prediction error of MAP over the future horizon and at steady state.

The control objective is given as follows:

$$J_{LRPC} = \sum_{j=n1}^{n2} [\hat{y}(t+j) - w(t+j)]^2 + \sum_{j=1}^{nu} \gamma(j) [\hat{y}(s|t+j-1) - w(s)]^2 \qquad (7.2)$$

where $n1$ is the minimum output prediction horizon, $n2$ is the maximum output prediction horizon, nu is the control horizon, $\gamma(j)$ is a steady-state error weighting sequence, and s denotes a value at steady-state.

The method of obtaining such model is discussed in the next section. The incorporation of the steady-state weighting term allows the controller to "look" not only at the next few prediction instances, but also further ahead to the steady state. It is similar to a cautious driver who is prepared for difficult driving conditions in the immediate vicinity but can also clearly look as far away as the final destination.

7.2.2 Identification

The control system is made adaptive by on-line estimation the coefficients in polynomials $\hat{A}(q^{-1})$ and $\hat{B}(q^{-1})$. In all previous attempts of predictive and adaptive blood pressure control, the model coefficients were basically obtained by a least squares(LS) approach:

$$J_{LS} = \frac{1}{T} \sum_{t=1}^{T} [y(t) - \hat{y}(t|t-k)]^2 \qquad (7.3)$$

which minimizes the total variance of past prediction errors. When a fixed time delay k is given, the LS approach produces a model describing the relationship between the MAP outputs and k^{th}-unit-delayed SNP infusions. The drawback in this approach lies in the fact that the model only provides an optimal k^{th}-step ahead prediction value in the least squares sense. Any further predictions beyond the k^{th}-step value would contain propagated prediction errors which might result in poor performance in a long-range predictive controller. Therefore, the following control-relevant identification strategy for long-range predictions was used to provide the best match for a trajectory of MAP responses:

$$J_{LRPI} = \frac{1}{T} \sum_{t=1}^{T} \sum_{j=n1}^{n2} [y(t) - \hat{y}(t|t-j)]^2 \tag{7.4}$$

The implementation of an estimator which satisfies eqn. 7.4 requires non-linear optimization. However, an elegant technique developed by Shook *et al.* (1991) as an alternative to the objective in 7.4 was used in this study to identify a MAP response model. It equates the power spectral density of eqn. 7.3 with a data prefilter $L(q^{-1})$ to that of eqn. 7.4. The prefilter is found on-line by using a spectral factorization routine. Then a LS-based estimator coupled with the $L(q^{-1})$ data prefilter is equivalent to the long-range predictive identification objective.

7.3 System Description

7.3.1 Equipment

The closed-loop control system used in this study is similar to the one described earlier in Chapter 6. The major equipment includes a Grass (Model 7D) monitoring system, an IMED[2] 929 computer-enabled drug infusion pump, and an IBM Model-70 PS/2 personal computer. Blood pressure was measured by a transducer and processed by the monitoring system which provides an analog MAP signal. The multitasking computer system acted as an operator station responsible for data acquisition and control calculation, as well as information display and front-panel command server. SNP as a manipulating variable was delivered by the IMED pump while other drugs for induced disturbances were infused by a Harvard[3] Model 975 pumping unit.

---

[2] IMED Corporation, San Diego, California

[3] Harvard Infusion Pump, South Natick, MA

7.3.2 Control Program

The overall control program was written in "C"-language and executed under QNX which is a realtime, multitasking operating system. The program consists of a number of tasks managed by a multi-purpose control system package (MULTICON) (Qiu *et al.*, 1988). The major tasks include a database manager, a timing scheduler, drug infusion pump driver, MAP signal receiver, an identification task, and a control task. The results from both identification and control tasks were processed by two validation modules. A set of first order models with different time delays were being regressed simultaneously in the identification task. The integrity of the models was analyzed to obtain the best model in the model validation module according to the following criteria:

- overall model gain is negative,
- overall model gain is within acceptable limits,
- model is stable,
- model produces the least prediction error.

Similarly the control signal validation module checks for any violation of infusion constraints and possible overdose. Should the MAP fall below a certain critical user-specified limit, the routine would reduce SNP infusion in an aggressive manner.

7.4 Experimental Studies

7.4.1 Methods

This study was approved by the Health Sciences Animal Care Committee of the University of Alberta. Experiments were performed on six healthy mongrel dogs of either gender weighing between 16 to 18 kg. Fig. 7.1 shows the schematic diagram of the whole experimental setup.

Surgical Preparation

Anesthesia was induced with sodium pentobarbital (30 mg kg^{-1}) and body temperature was maintained between 37°C and 38°C by a heating pad. Ventilation with an oxygen (O_2) enriched air mixture to maintain arterial O_2 tension greater than 120 mmHg was provided by a respirator (Harvard Apparatus, South Natick, MA) following intubation with a cuffed endotracheal tube. Anesthesia was maintained by a constant infusion (4 ml kg^{-1} hr^{-1}) of pentobarbital (3 mg kg^{-1} hr^{-1}) in 0.9 % NaCl. The right external jugular vein was cannulated for fluid loading (10 ml kg^{-1} of 6 % Dextran 70 in 0.9 % NaCl) and anesthetic infusion. A triple lumen catheter was inserted in the left femoral vein for infusion of vasoactive drugs. A catheter was placed in the aortic arch via the left femoral artery to measure aortic pressure and allow blood sampling at regular intervals for determination of blood gases, pH and electrolytes. Normal carbon dioxide tension was maintained by continuous monitoring of end-tidal CO_2 (LB2 analyzer[4]). Acidosis and hypokalaemia were corrected, when required, with sodium bicarbonate and potassium chloride, respectively. ECG tracings, heart rate derived from R-R intervals, systolic, diastolic and MAP were continuously recorded on a Grass (Model 7D) polygraph recorder.

Challenge Protocol

Each study began with a controller initialization sequence which consisted of the infusion of SNP (7 to 15 μg $kg^{-1}min^{-1}$) to produce a small hypotensive response. This also constituted the initial probing sequence for the on-line estimation scheme.

---

[4] Beckmann Instruments, Fullerton, CA

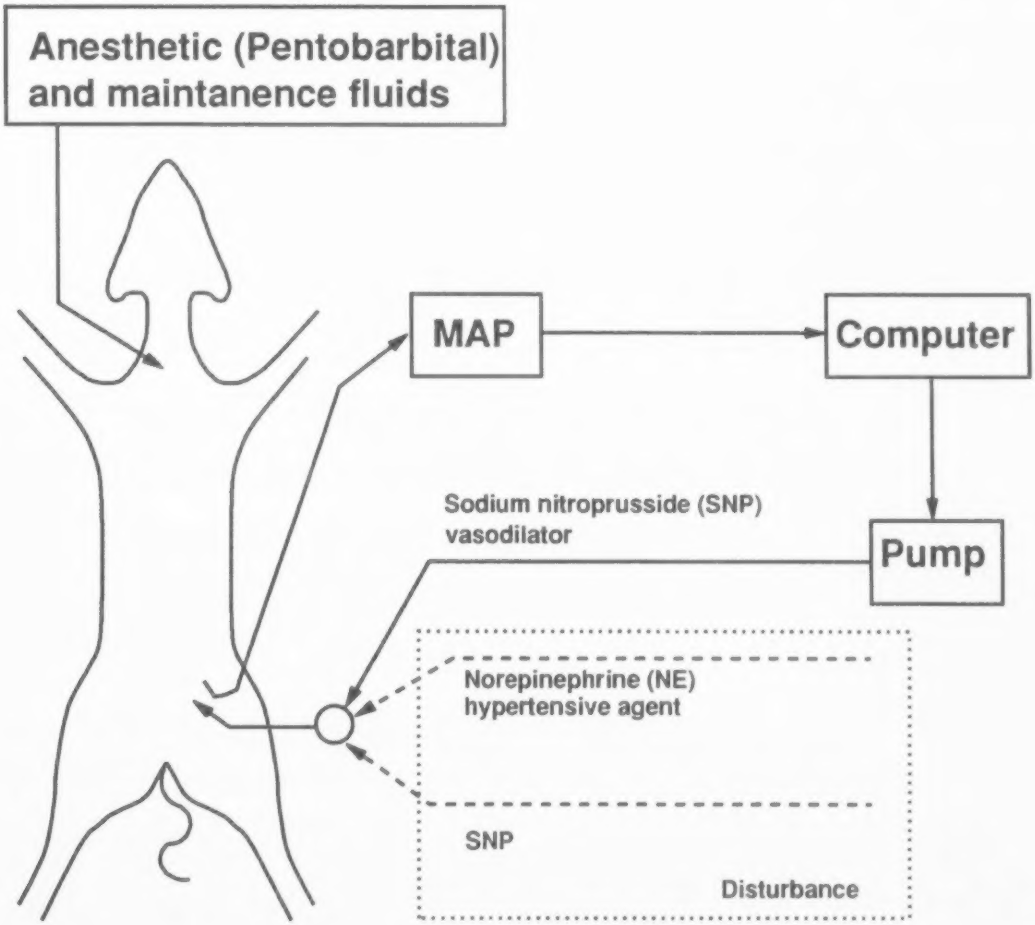
Anesthetic (Pentobarbital)
and maintanence fluids
MAP
Computer
Sodium nitroprusside (SNP)
vasodilator
Pump
Norepinephrine (NE)
hypertensive agent
SNP
Disturbance

Figure 7.1: Schematic diagram of the experimental setup

Thereafter, the performance of the controller was evaluated for setpoint changes by requesting a computer controlled setpoint changes of -30 mmHg from the original baseline. Controller performance was also tested during a series of unpredictable disturbances produced by the administration of vasoactive drugs (see below). Two maintained disturbances were made consecutively, one in the absence and the other in the presence of closed-loop computer control. Deviations from baseline MAP under these two conditions were compared. In order to ensure that the disturbances in MAP were made from equivalent baseline pressures, MAP was lowered to similar levels by either a constant infusion of SNP (absence of closed-loop control) or by the controller (presence of closed-loop control). Transient disturbances were also introduced by vasoactive drug injections.

Vasoactive Drug Administration

SNP was administered by the computer-controlled delivery system to reduce MAP. Concentrations of 200 μg ml^{-1} or 400 μg ml^{-1} were infused depending on the initial sensitivity of each animal. Maintained and transient disturbances in MAP were induced by either 10 min continuous intravenous infusions (Harvard Infusion Pump, Model 975, South Natick, MA) or by bolus doses, respectively, of vasoactive drugs. Hypertensive disturbances were elicited by either an intravenous infusion (0.5 to 4 μg kg^{-1} min^{-1}) or bolus administration (2.5 to 5.5 μg kg^{-1}) of norepinephrine (NE). Hypotensive disturbances were produced by either an intravenous infusion (6 to 32 μg kg^{-1} min^{-1}) or bolus administration (9 to 32 μg kg^{-1}) of SNP. NE infusion rates were chosen to cause approximately a 30 mmHg deviation in MAP, whereas SNP infusion rates were selected to be 75 % of the constant SNP infusion rate in the absence of computer control. Bolus doses of NE and SNP were chosen to produce transient changes in MAP of 30 and 10 mmHg, respectively. Recovery periods of at least 10 min were allowed between maintained and transient disturbances.

7.4.2 Results

Five sets of performance response are plotted in figs. 7.2 to 7.6, corresponding respectively to setpoint tracking after initialization, NE infusion, SNP infusion, NE bolus, and SNP bolus. Each figure shows the average deviation of MAP from the target with an error bar of one standard deviation. The average value was calculated by taking the mean of MAP measurements from the same runs on different subjects at 30-second intervals for fig. 7.2 and one-minute intervals for the rest of the figures. A typical trajectory of SNP infusions (dashed line) is also plotted along with the average trajectory (solid line) in fig. 7.2. The uncontrolled results are plotted above the computer-controlled ones for comparison in figs. 7.3 and 7.4. The average time for the probing signal is 2.88 ±0.54 min (mean ± standard mean error). The average time for MAP to fall into the acceptable region (± 5 mmHg of the target) after the probing signal is 2.44 ±0.31 min. No overshoot beyond 5 mmHg is observed during all setpoint changes.

During the time of NE infusion with no corrective control actions (*i.e.* infusion of SNP was kept constant), MAP increased to an average deviation of 43.88 ± 4.00 mmHg after ten minutes of infusion. The recovery period shown in fig. 7.3 clearly took more than ten minutes. The response with the control system shows a drastic improvement in which average MAP deviation climbed to only 8.69 ± 1.21 mmHg. The average time outside the acceptable region is reduced to only 2.13 ± 0.60 min.

In fig. 7.4, the average deviation of MAP is not as large as that caused by NE, the reason being that the disturbance SNP infusion was calculated as a percentage of

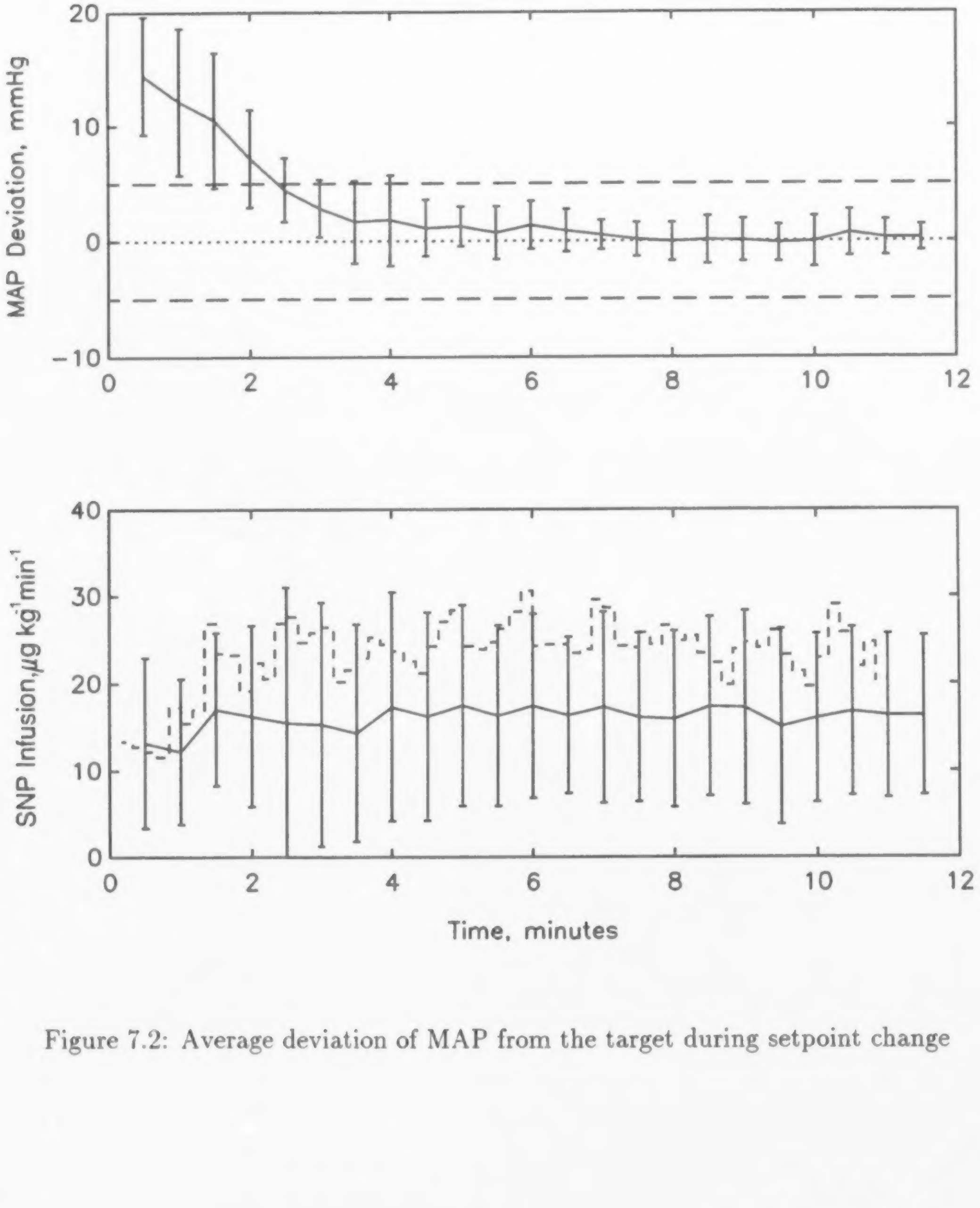

Figure 7.2: Average deviation of MAP from the target during setpoint change

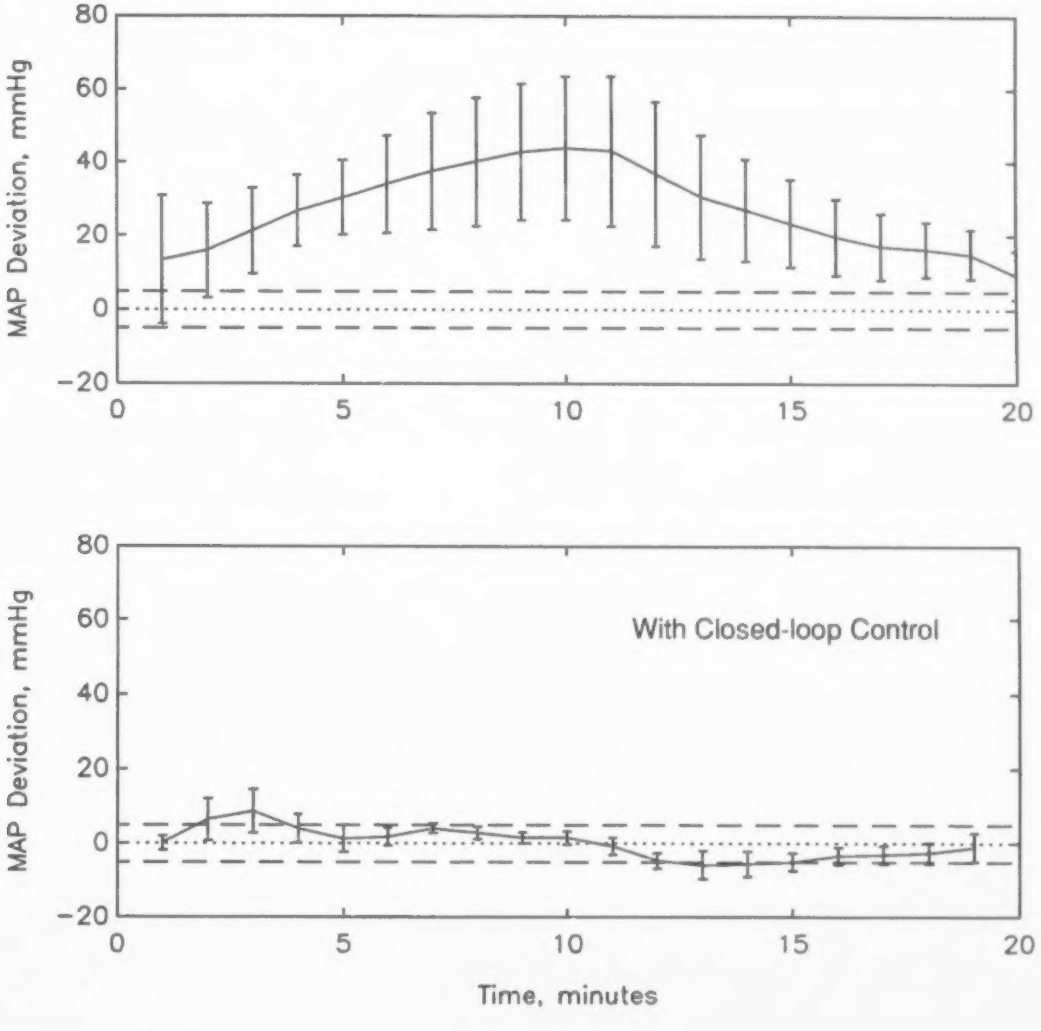

Figure 7.3: Average deviation of MAP from the target during NE infusion

the computer-controlled SNP infusion. Larger doses of disturbance SNP could have been used, but would have caused so much hypotension that no computer-controlled SNP infusion was required. Without computer control but at a constant infusion of disturbance SNP, average MAP deviates from the setpoint by 5.93 ± 0.36 mmHg. With computer control, the MAP was maintained within the acceptable region during the whole disturbance SNP infusion challenge.

In response to NE bolus (see fig. 7.5), the average MAP under computer control made a deviation of 14.70 ± 1.57 mmHg and was outside the acceptable region for 1.75 ± 0.14 min. An average maximum overshoot of 5.44 ± 1.11 mmHg was recorded at t=7 min. In fig. 7.6, SNP bolus made an average deviations of 12.61 ± 1.74 mmHg for computer-controlled responses. The control system brought the MAP within the region in 2.67 ± 0.16 min. after the injection with an overshoot of 4.86 ± 1.45 mmHg. The control system maintained stability throughout all challenges.

7.5 Discussion

The first task of the MAP controller was to induce hypotension in an efficient and robust manner. During the time of induction in fig. 7.2, MAP was brought to the setpoint without overshoot. The average trajectory of SNP infusion shows no sign of ringing or oscillation which were present in most of the one-step ahead control results by others. The error bars measuring one standard deviation also indicate that the variations of MAP were well within the acceptable region. Only 12 out of 312 MAP samples collected in a six-minute interval after four minutes of setpoint change were beyond ±5 mmHg from the target. It means that over 96 % of MAP regulation is within the target region.

The method for assessing the performance of this controller in the presence

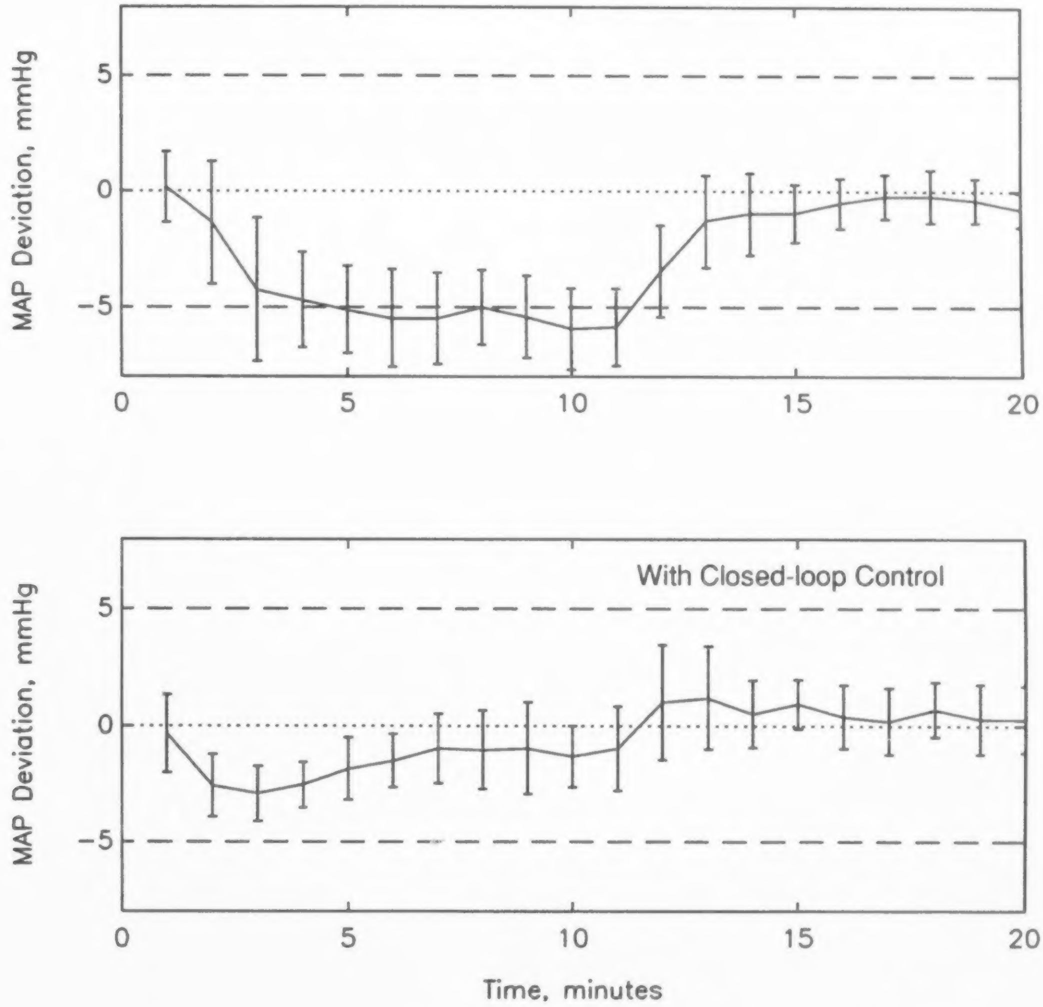

Figure 7.4: Average deviation of MAP from the target during SNP infusion

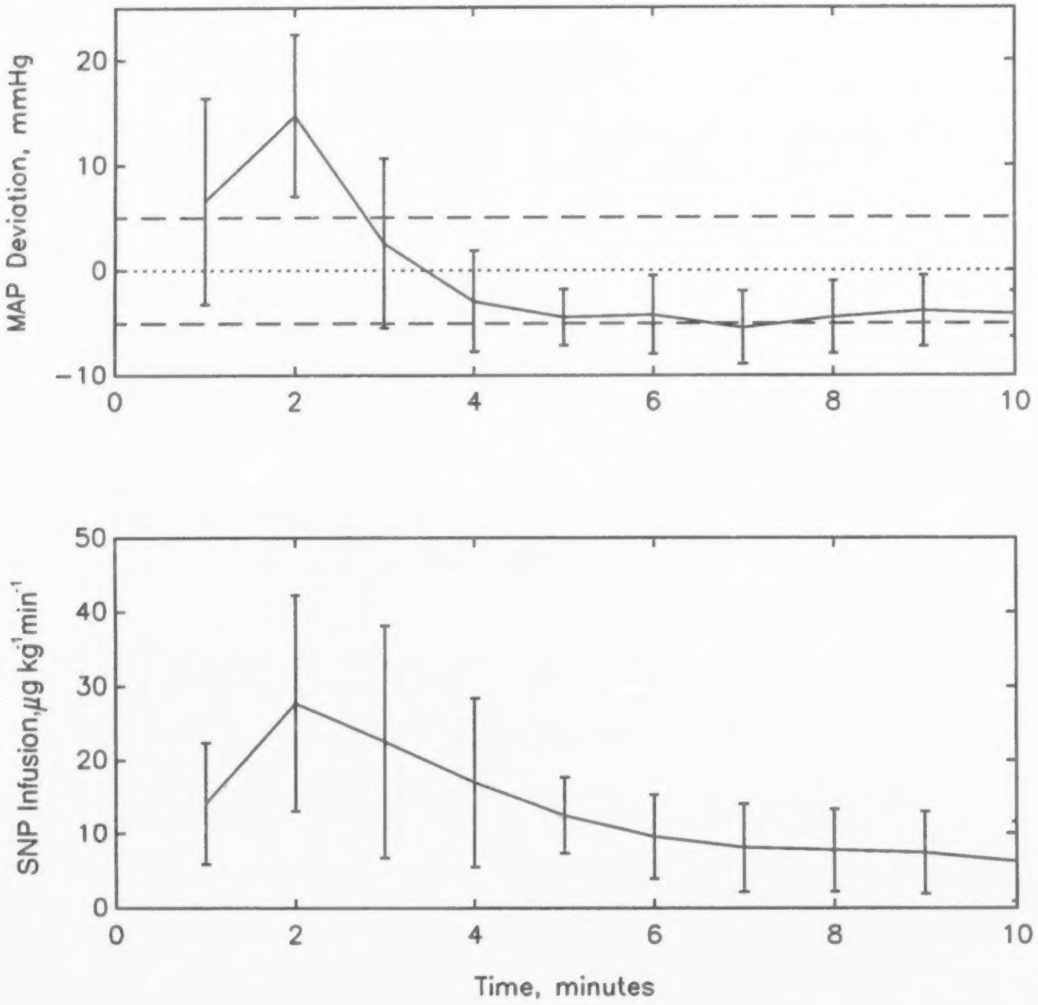

Figure 7.5: Average deviation of MAP from the target during NE injection

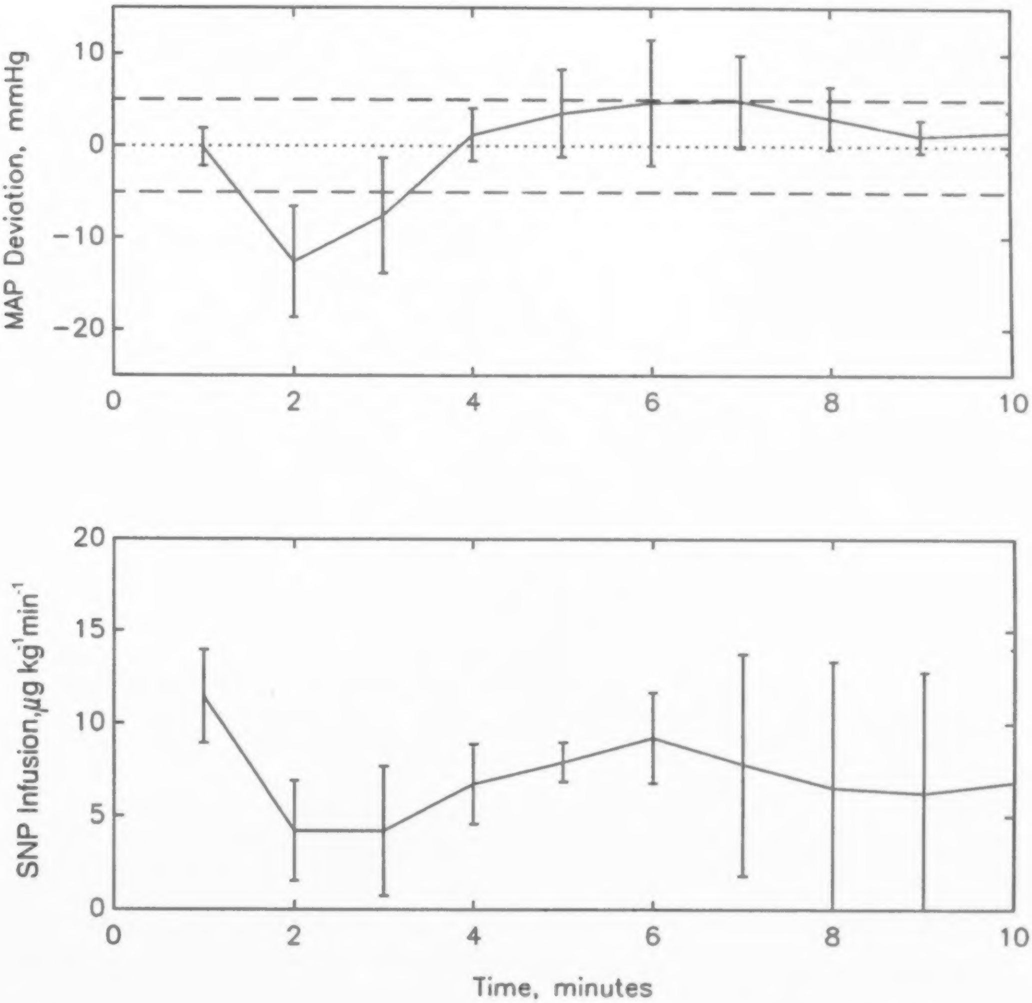

Figure 7.6: Average deviation of MAP from the target during SNP injection

of disturbances is significantly different from other previous studies described in the introductory section. The challenge protocol which asked for two disturbances of the same dose to be introduced consecutively with and without computer control was designed to ensure that unpredictable disturbances indeed produced a noticeable effect on the subject before the controller responded. It should be noted that without computer control, the autonomic mechanism of blood pressure regulation was still in operation in addition to external drug infusions. Therefore, "uncontrolled" or open-loop MAP responses implies autonomic body regulation only. The continuous infusion of NE and SNP as disturbances simulated a slow change in the subject's sensitivity to drug therapy. Infusion of SNP as a disturbance to lower MAP simulated an increase in sensitivity of the subject to the computer-controlled SNP infusion, whereas infusing NE implied a decrease in sensitivity. In fig. 7.3, NE infusion had an enormous impact on raising the MAP from the target while no corrective response was taken by the controller. It simulates the situation where a patient's autonomic regulatory system fails to maintain the blood pressure at a normal level especially after cardiac surgery. When the same NE disturbance was introduced, the controller responded by increasing computer-controlled SNP infusion to keep the MAP close to the target. After NE infusion was stopped, MAP started to drop to 5.66 ± 0.78 because of the fading NE effect and high SNP infusion. In fig. 7.4, the 75 % selection approach for constant SNP infusion did not produce as large and drastic a drop as that by NE, but it allowed the transient performance of the controller to be demonstrated. Making a large MAP drop to challenge the controller of course would be more desirable. However, the controller in most of the cases would turn off computer-controlled SNP infusion and wait for the MAP to rise. In such a case, the rate of returning MAP to target does not depend on the controller, but on the subject's metabolic rate.

Injections of SNP and NE in high dosage were used to test the controller for its ability to cope with sudden changes in MAP. Unlike chemical processes where some kinds of disturbances are measurable and hence feedforward control is possible, physiological disturbance or upsets are usually unforeseeable. In this situation, corrective action cannot be taken until after the disturbance appears in the blood pressure signal. It implies that the large and abrupt deviations of MAP shown in both figs. 7.5 and 7.6 were unavoidable because the control system did not "know" of the disturbance drug injection until the MAP started to change. Nevertheless, the control system managed to return MAP to the target by rapidly changing SNP infusions. Since the disturbance effect by injection did not last as long as that by infusion, the fast return of MAP shown in both figures was followed by overshoot. It should be noted that the control system or even an anesthesiologist would not be able to know *a priori* the duration of the disturbance effect, and hence such overshoots as shown in figs. 7.5 and 7.6 were indeed inevitable.

7.6 Conclusions

A long-range predictive adaptive control system for computerized drug delivery has been evaluated for the control of mean arterial pressure on six mongrel dogs. This truly model-based adaptive control system is a combination of generalized predictive control with steady-state error weighting and long-range predictive identification. It has been tested for setpoint tracking and in the presence of a set of unpredictable infusions and injections of nitroprusside and norepinephrine. It is found that an average of 2.44 ± 0.31 min after probing was required to reach the hypotension target. No overshoot nor ringing was followed after setpoint change. 96.2 % of the MAP regulation was within ± 5 mmHg of the target. In spite of unpredictable infusion disturbances which would have caused a maximum average of 43.88 mmHg deviation, the control system brought the mean arterial pressure to the target in about two minutes. The system also responded favorably during sudden upsets of pressure when, in fact, the magnitude of the upsets were unavoidable.

Bibliography

Clarke, D.W., C. Mohtadi, and P.S. Tuffs. Generalised predictive control – Part I. The basic algorithm. *Automatica*, 23(2):137–148, 1987.

Colvin, J.R. and G.N.C. Kenny. Automatic control of arterial pressure after cardiac surgery. *Anaesthesia*, 44:37–41, 1989.

Cosgrove III, D.M., J.H. Petre, J.L. Waller, J.V. Roth, C. Shepherd, and L.H. Cohn. Automated control of postoperative hypertension: A prospective, randomized multicenter study. *Ann. Thorac. Surg.*, 47:678–683, 1989.

de Asla, R.A., A.M. Benis, R.A. Jurado, and R.S. Litwak. Management of postcardiotomy hypertension by microcomputer-controlled administration of sodium nitroprusside. *J. Thorac. Cardiovasc. Surg.*, 89:115–120, 1985.

Katona, P.G. Automated control of physiological variables and clinical therapy. *CRC Critical Reviews in Biomedical Engineering*, 8(4):281–310, 1982.

Kaufman, H., R. Roy, and X. Xu. Model reference adaptive control of drug infusion rate. *Automatica*, 20(2):205–209, 1984.

Linkens, D.A. Computer control in biomedicine. In S. Bennett and D.A. Linkens, editors, *Real-Time Computer Control*, chapter 14. P. Peregrinis, 1984.

Linkens, D.A. and S.S. Hacisalihzade. Computer control systems and pharmalogical drug administration: a survey. *J. Med. Eng. & Tech.*, 14(2):41–54, 1990.

Linkens, D.A., M. Mahfouf, and A.J. Asbury. Multivariable generalized predictive control for anaesthesia. In *Proc. European Control Conference*, pages 1630–1635, Grenoble, France, 1991.

Meline, L.J., D.R. Westenskow, N.L. Pace, and M.N. Bodily. Computer-controlled regulation of sodium nitroprusside infusion. *Anesth. Analg.*, 64:38–42, 1985.

Meline, L.J., D.R. Westenskow, A. Somerville, R.T. Wernick, J. Jacobs, and N.L. Pace. Evaluation of two adaptive sodium nitroprusside control algorithms. *J. Clin. Monit.*, 2:79–86, 1986.

Millard, R.K., P. Hutton, E. Pereira, and C. Prys-roberts. On using a self-tuning controller for blood pressure regulation during surgery in man. *Comput. Biol. Med.*, 17(1):1–18, 1987.

Millard, R.K., C.R. Monk, and C. Prys-Roberts. Surgery using a volatile anaesthetic. *IEE Proc.-D*, 135(2):95–105, March 1988.

Qiu, Z., M. Fortuna, and D.G. Fisher. Multi-purpose control system package. Technical report, Dept. of Chem. Eng., University of Alberta, 1988.

Reid, J.A. and G.N. Kenny. Evaluation of closed-loop control of arterial pressure after cardiopulmonary bypass. *Br. J. Anaesth.*, 59:247–255, 1987.

Sheppard, L.C., N.T. Kouchoukos, J.F. Shotts, and F.D. Wallace. Regulation of mean arterial pressure by computer control of vasoactive agents in postoperative patients. In *Computers in Cardiology*, pages 91–94, Rotterdam, Netherlands, October 2–4 1975.

Shook, D.S., C. Mohtadi, and S.L. Shah. Identification for long range predictive control. *IEE Proc.-D*, 138(1):75–84, 1991.

Slate, J.B. and L.C. Sheppard. Automatic control of blood pressure by drug infusion. *IEE Proc.-A*, 129(9), December 1982a.

Slate, J.B. and L.C. Sheppard. A model-based adaptive blood pressure controller. In *IFAC Identification and System Parameter Estimation*, pages 1437–1442, 1982b.

Slate, J.B., L.C. Sheppard, V.C. Rideout, and E.H. Blackstone. Closed-loop nitroprusside infusion: modeling and control theory for clinical application. In *IEEE 1980 International Symposium on Circuits and Systems*, pages 482–488, 1980.

Stern, K.S., H.J. Chizeck, B.K. Walker, P.S. Krishnaprasad, P.J. Dauchot, and P.G. Katona. The self-tuning controller: comparison with human performance in the control of arterial pressure. *Annals of Biomedical Engineering*, 13:341–357, 1985.

Voss, G.I., H.J. Chizeck, and P.G. Katona. Self-tuning controller for drug delivery systems. *Int. J. Cont.*, 47(5):1507–1520, May 1988.

Westenskow, D.R., L. Meline, and N.L. Pace. Controlled hypotension with sodium nitroprusside: anesthesiologist vesus computer. *J. Clin. Monit.*, 3:80–86, 1987.

Yu, C., R.J. Roy, H. Kaufman, and B.W. Bequette. Multiple-model adaptive predictive control of mean arterial pressure and cardiac output. Submitted to IEEE Trans on Biomedical Engineering, 1991.

Chapter 8

Conclusions

8.1 General Discussions and Conclusions

This thesis has dealt with two related areas of research; first, the incorporation of a terminal matching condition in both LRPC and LRPI formulation; second, the development of both software and hardware for a closed-loop drug delivery system, and the evaluation of such a system for the regulation of mean arterial pressure. A synopsis of all conclusions and contributions is given as follows:

1. Terminal Condition in LRPC

The terminal matching condition was first introduced into LRPC as a weighting on the steady-state error. The LRPC control law in this newly modified form minimizes the squares of prediction errors over a future prediction horizon and at steady-state. The incorporation of such weighting in the control law naturally requires the prediction of the steady-state output. For the ARIMAX model, a direct method for calculating such a prediction was proposed so that further iterations of the Diophantine identity were not necessary. With a convolution model, a complete set of dynamic response coefficients was no longer required. Only a few step response coefficients plus the steady-state gain were used to approximate the major dynamics of a process. These modeling results pave the way for the derivation of an adaptive long-range predictive controller with reduced numerical computational requirements.

Closed-loop analyses of the steady-state error weighting with Generalized Predictive Control (GPC) (Clarke *et al.*, 1987), a popular version of LRPC, show that the combination of a relatively short prediction horizon and proper selection of the steady-state error weighting was comparable to using a large prediction horizon. This weighting term is also shown to be more advantageous than ordinary control weighting. Most interestingly, the steady-state error weighting not only causes zero offset at steady-state even when a model gain was in error, but also results in a characteristic equation with **one** less order, *i.e.* it eliminated the pole introduced by the integrator from the disturbance modeling term.

The robustness introduced by the steady-state error weighting was demonstrated by the fact that it provided a stabilizing effect for open-loop stable systems even in the presence of modeling error and non-minimum zeros. Two strategies were also included to serve as guidelines for selecting such a weighting with other primary LRPC tuning parameters at the commissioning stage.

2. Terminal Condition in LRPI

Shook *et al.* (1991) proposed a new, control-relevant identification algorithm for LRPC via spectral factorization of a large $(2\cdot(n2-1))$ polynomial where $n2$ was the output prediction horizon. This algorithm produced a data pre-filter $L(q^{-1})$ which was of order $n2-1$. The resulting model could give the best predictions of the first $n2$ dynamic responses in the least squares sense (Shook, 1991). For a prediction horizon $n2$ approaching infinity, Shook's algorithm was theoretically shown in Chapter 4 to change the ARMAX model scheme from an equation error scheme to an output error one. However, this convergence property was not valid for the ARIMAX model because $L(q^{-1})$ with ARIMAX did not converge to a finite polynomial. Practically speaking, for a large prediction horizon, the algorithm suffers from heavy computation and poor convergence. This disadvantage implies that a relatively short prediction horizon is beneficial to both LRPC and LRPI because of further reduction in the computational load of adaptive controllers.

Although a terminal matching condition in the form of Shook's LRPI algorithm was not feasible, analyzing the derivation of the LRPI algorithm revealed the fact that the condition could be indirectly realized by identification of the process gain. A simple algorithm similar to the non-minimal model predictor of Lu and Fisher (1990) was, thus, proposed for on-line estimation of a process gain. As a result, a multi-step, adaptive, predictive controller including the terminal condition was synthesized from the combination of a long-range predictive control law such as GPC, the steady-state error weighting in Chapters 2 and 3, Shook's LRPI algorithm, and an on-line gain estimation algorithm such as the one in Section 4.3.2. The performance of using this controller in simulation studies show that accurate estimation of process gain combined with steady-state error weighting improve the overall robustness of the controller even with a very small prediction horizon and large model-plant mismatch.

3. Development of a LRPC System for MAP Regulation

The survey in Chapter 5 shows many examples of applying modern control engineering to mean arterial pressure (MAP) regulation. The milestones marked by Sheppard and Slate's pioneering work (Sheppard *et al.*, 1975; Sheppard and Sayers, 1977; Slate and Sheppard, 1982) and the IVAC[1] Titrator (Cosgrove III *et al.*, 1989) have encouraged further research into closed-loop control drug delivery systems based on more advanced control strategies among which GPC, a widely used version of LRPC, has been claimed to possess the best features of many earlier algorithms (Mohtadi, 1987).

---

[1] IVAC Corporation, San Diego, California

Software was developed for implementing adaptive GPC in a personal computer running under QNX, a realtime operating system[2] with multitasking and networking capabilities. It provided on-line identification of models using recursive least squares, on-line changes of the output prediction horizon ($n1$ and $n2$), control horizon (nu), setpoint pre-filter($R(q^{-1})$), and output filter ($P(q^{-1})$). All software programs were written in the "C"-language.

A closed-loop LRPC system for MAP regulation was developed by interfacing the computer with a computerized drug infusion pump and a analog-to-digital converter receiving the MAP signal from a patient monitoring device. A preliminary test shows the effectiveness of both unconstrained and constrained GPC on MAP regulation. Problems caused by model-plant mismatch, varying time delays and non-linearity were all handled satisfactorily by the system (See Chapter 6). The constrained GPC version was developed by Mutha (1991).

4. Evaluation of the System for MAP regulation

GPC with steady-state error weighting and Shook's (1991) LRPI algorithm was implemented in the final version of the control system. The performance of the system was evaluated for setpoint tracking and in the presence of a set of infusions and injections of nitroprusside and norepinephrine. The results once again demonstrated the excellent control and the effectiveness as well as the robustness of the system for MAP regulation. Although this system was applied to blood pressure regulation, control of other physiological parameters was expected to be effective as long as frequent sampling of the physiological signal was possible.

---

[2] QNX, Quantum Software Systems, Ltd., 1988

8.2 Recommendations

1. Accurate identification of process gain is imperative for improving robustness. Further research in other on-line gain estimation techniques is desirable.
2. Expert system approach has been considered for many control areas where parametric models are not necessary or cannot be realized. In blood pressure control, a supervisory knowledge-based expert system could look after the initial control period, detect artifacts and, and perform minor tuning.
3. Once a supervisory MAP control is available, performance evaluation should proceed to clinical trials in an ethics-approved manner.
4. Several attempts have been initiated to simultaneously control more than one physiological parameters such as MAP plus cardiac output. However, the lack of a clinically acceptable method of measuring cardiac output remains a hindrance to further study of multiple drug infusion control systems. A possible approach for accessing the parameters is by means of a "software sensor" which infers the primary parameters from other readily available physiological signals such as pulse rate and the blood pressure wave (*i.e.* an inferential control system).

Bibliography

Clarke, D.W., C. Mohtadi, and P.S. Tuffs. Generalised predictive control – Part I. The basic algorithm. *Automatica*, 23(2):137–148, 1987.

Cosgrove III, D.M., J.H. Petre, J.L. Waller, J.V. Roth, C. Shepherd, and L.H. Cohn. Automated control of postoperative hypertension: A prospective, randomized multicenter study. *Ann. Thorac. Surg.*, 47:678–683, 1989.

Lu, W. and D.G. Fisher. Nonminimal model based long range predictive control. In *Proc. American Control Conference*, pages 1607–1613, San Diego, CA, 1990.

Mohtadi, C. *Advanced Self-Tuning Algorithms.* D.Phil thesis, University of Oxford, 1987.

Mutha, R.K. Constrained long range predictive control. M.Sc. thesis, University of Alberta, 1990.

Sheppard, L.C., N.T. Kouchoukos, J.F. Shotts, and F.D. Wallace. Regulation of mean arterial pressure by computer control of vasoactive agents in postoperative patients. In *Computers in Cardiology*, pages 91–94, Rotterdam, Netherlands, October 2–4 1975.

Sheppard, L.C. and B.McA. Sayers. Dynamic analysis of the blood pressure response to hypotensive agents studied in post surgical patients. *Comput. Biomed. Res.*, 10:237–246, 1977.

Shook, D.S. *Identification Issues in Long-Range Predictive Control.* Ph.D. thesis, University of Alberta, 1991.

Shook, D.S., C. Mohtadi, and S.L. Shah. Identification for long range predictive control. *IEE Proc.-D*, 138(1):75–84, 1991.

Slate, J.B. and L.C. Sheppard. A model-based adaptive blood pressure controller. In *IFAC Identification and System Parameter Estimation*, pages 1437–1442, 1982.

Appendix A

Example: Simultaneous Offset Removal and Integrator-Pole Elimination by Steady-State Error Weighting

Without any loss of generality, a simple open-loop stable process is represented by the following time series model:

$$Ay(t) = Bu(t-1) + \frac{\xi(t)}{\Delta} \qquad (A.1)$$

The general vector form of control law from GPC with steady-state error weighting is given as follows:

$$\mathbf{U} = \left[\mathbf{G}^T\mathbf{G} + \Lambda + \mathbf{G_s}^T\Gamma\mathbf{G_s}\right]^{-1}\left[\mathbf{G}^T(\mathbf{W}-\mathbf{f}) + \mathbf{G_s}^T\Gamma(\mathbf{W}_s - \mathbf{f_s})\right] \qquad (A.2)$$

The dynamic matrix coefficients in **G** are equivalent to the step response coefficients obtained via polynomial division,

$$\frac{B}{A\Delta} = g_0 + g_1q^{-1} + g_2q^{-2} + \cdots\cdots \qquad (A.3)$$

The steady-state gain of the process is defined as g_s such that

$$\frac{B(1)}{A(1)} = g_s \qquad (A.4)$$

A one-step ahead control with steady-state error weighting is configured by choosing the following settings:

$$\left.\begin{array}{rcl} n1 &=& 1 \\ n2 &=& 1 \\ nu &=& 1 \\ \lambda &=& 0 \\ \gamma &=& \gamma' g_0 \end{array}\right\} \tag{A.5}$$

The peculiar setting of γ will make the sequel more readable. With this configuration, the matrices in eqn. A.2 are reduced to scalar quantities:

$$\mathbf{G} = g_0$$

$$\mathbf{G}_s = g_s$$

Then the control law can be expressed as an algebraic equation:

$$\begin{aligned} \Delta u(t) &= \frac{g_0}{g_0^2 + g_s^2 g_0 \gamma'}[w(t) - F_1 y(t) - H_1 \Delta u(t-1)] \\ &\quad + \frac{g_s g_0 \gamma'}{g_0^2 + g_s^2 g_0 \gamma'}[w(t) - F_s y(t) - H_s \Delta u(t-1)] \end{aligned} \tag{A.6}$$

After certain re-arrangement, eqn. A.6 is written below in a linear form:

$$R\Delta u(t) = Vw(t) - Sy(t) \tag{A.7}$$

where

$$R = g_0 + q^{-1}H_1 + g_s\gamma'(g_s + q^{-1}H_s) \tag{A.8}$$

$$V = 1 + g_s\gamma' \tag{A.9}$$

$$S = F_1 + g_s\gamma' F_s \tag{A.10}$$

Before proceeding with the closed-loop transfer function derivation, one needs to examine the Diophantine identities which are essential for model prediction from eqn. A.1

$$1 = A\Delta E_j + q^{-j}F_j \tag{A.11}$$

$$BE_j = G_j + q^{-j}H_j \tag{A.12}$$

Since polynomials A are E_j monic, for $j = 1$,

$$\begin{aligned} 1 &= A\Delta + q^{-1}F_1 && \text{(A.13)} \\ B &= g_0 + q^{-1}H_1 && \text{(A.14)} \end{aligned}$$

The two identities in A.11 and A.12 can be combined into one identity as follows:

$$\begin{aligned} \frac{B}{A\Delta} &= BE_j + q^{-j}\frac{F_jB}{A\Delta} \\ &= G_j + q^{-j}\left[H_j + \frac{F_jB}{A\Delta}\right] \\ \left[\frac{B}{A\Delta} - G_j\right] &= \left[H_j + \frac{F_jB}{A\Delta}\right]q^{-j} && \text{(A.15)} \end{aligned}$$

For sufficiently large j,

$$\begin{aligned} \left[\frac{B}{A\Delta} - G_j\right] &\approx g_sq^{-j} + g_sq^{-j-1} + g_sq^{-j-2} + \cdots \\ &\approx \frac{g_s}{\Delta}q^{-j} && \text{(A.16)} \end{aligned}$$

In the limit as $j \to \infty$,

$$\left[H_j + \frac{F_jB}{A\Delta}\right] = \frac{g_s}{\Delta} \tag{A.17}$$

So far, the equations necessary for closed-loop transfer function derivation have been examined. Now consider the general form of a closed-loop transfer function derived from the model in eqn. A.1 and the linear control law in eqn. A.7:

$$y(t) = \frac{BVq^{-1}w(t) + R\Delta x(t)}{RA\Delta + q^{-1}BS} \tag{A.18}$$

Substituting eqns. A.8 and A.10 into the characteristic polynomial gives

$$\mathcal{C} = \underbrace{(g_0 + q^{-1}H_1)A\Delta}_{1} + g_s\gamma'(g_s + q^{-1}H_s)A\Delta + \underbrace{q^{-1}BF_1}_{2} + q^{-1}g_s\gamma'BF_s \tag{A.19}$$

After substituting eqns. A.14 and A.13 into terms 1 and 2 respectively, the summation of the two terms becomes B,

$$\mathcal{C} = B + g_s\gamma'(g_s + q^{-1}H_s)A\Delta + q^{-1}g_s\gamma'BF_s \tag{A.20}$$

$$= B + g_s\gamma' \left[g_s A\Delta + q^{-1} A\Delta \underbrace{\left(H_s + \frac{BF_s}{A\Delta} \right)}_{3} \right] \tag{A.21}$$

Eqn. A.17 shows that term 3 can be further reduced as follows:

$$\begin{aligned} \mathcal{C} &= B + g_s\gamma' \left[g_s A\Delta + q^{-1} A g_s \right] \\ &= B + g_s^2 \gamma' A \end{aligned} \tag{A.22}$$

This result indicates that the additional order due to integral action is removed. The offset-free property is shown by examining the closed-loop transfer function in eqn. A.18 at steady state. Substituting eqns. A.22 and A.9 into the transfer function yields

$$y(t) = \frac{B(1 + g_s\gamma')q^{-1}w(t) + R\Delta x(t)}{B + g_s^2\gamma' A} \tag{A.23}$$

At steady state,

$$\begin{aligned} y(t) &= \frac{B(1)(1 + g_s\gamma')w(t)}{B(1) + g_s^2\gamma' A(1)} \\ &= \frac{g_s(1 + g_s\gamma')}{g_s + g_s^2\gamma'} w(t) \\ &= w(t) \end{aligned}$$

The above shows how simultaneous offset removal and order reduction are accomplished by steady-state error weighting. These properties are also true for $n2$ and nu greater than one. Although no rigorous and generalized proof is available (mainly because R, V, and S are too complicated to analyze for $n2, nu > 1$), similar derivations as above can be obtained for $n2$ and nu larger than one.